

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2004, 09:06:13 ; Search time 40 Seconds
(without alignments)
580.283 Million cell updates/sec

Title: US-10-089-688-2
Perfect score: 1781
Sequence: 1 MPFPFGSHKSPADIVKNL.....RVKPRTRGIRDLKRPAAQEA 350

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1614.5	90.7	341	3	US-09-190-965-3	Sequence 3, Appli
2	1614.5	90.7	341	3	US-09-470-253-3	Sequence 3, Appli
3	1297.5	72.9	337	3	US-09-190-965-1	Sequence 1, Appli
4	1297.5	72.9	337	3	US-09-470-253-1	Sequence 1, Appli
5	1117	62.7	339	3	US-09-190-965-4	Sequence 4, Appli
6	1117	62.7	339	3	US-09-470-253-4	Sequence 4, Appli
7	1054.5	59.2	377	3	US-09-190-965-5	Sequence 5, Appli
8	1054.5	59.2	377	3	US-09-470-253-5	Sequence 5, Appli
9	358.5	20.1	165	4	US-09-248-796A-14303	Sequence 14303, A
10	166	9.3	125	4	US-09-248-796A-15202	Sequence 15202, A
11	163	9.2	90	4	US-09-248-796A-15201	Sequence 15201, A
12	115.5	6.5	630	4	US-09-248-796A-20275	Sequence 20275, A
13	113	6.3	541	4	US-09-134-000C-5420	Sequence 5420, Ap
14	113	6.3	2184	4	US-09-417-485D-6	Sequence 6, Appli
15	110.5	6.2	533	3	US-09-134-001C-4053	Sequence 4053, Ap
16	106	6.0	3959	2	US-08-970-269A-30	Sequence 30, Appl
17	106	6.0	3959	2	US-09-407-562-30	Sequence 30, Appl
18	105.5	5.9	3878	4	US-09-914-259-11	Sequence 11, Appl
19	105	5.9	496	4	US-09-543-681A-6465	Sequence 6465, Ap
20	105	5.9	627	4	US-09-248-796A-20614	Sequence 20614, A
21	105	5.9	776	1	US-08-021-601-2	Sequence 2, Appli
22	105	5.9	776	1	US-08-082-849B-2	Sequence 2, Appli
23	105	5.9	776	5	PCT-US94-01624-2	Sequence 2, Appli
24	104	5.8	933	4	US-09-248-796A-20513	Sequence 20513, A
25	103.5	5.8	1388	3	US-09-572-191-2	Sequence 2, Appli
26	103.5	5.8	1388	3	US-09-723-262-2	Sequence 2, Appli
27	103.5	5.8	1388	3	US-09-723-219-2	Sequence 2, Appli

28	103.5	5.8	2474	4	US-08-265-967C-3	Sequence 3, Appli
29	103.5	5.8	2474	4	US-08-305-790B-4	Sequence 4, Appli
30	103	5.8	1155	4	US-09-543-681A-6286	Sequence 4, Appli
31	103	5.8	1388	2	US-08-685-576-4	Sequence 4, Appli
32	101.5	5.7	467	4	US-09-543-681A-4262	Sequence 4262, Ap
33	101	5.7	489	4	US-09-248-796A-19030	Sequence 19030, A
34	101	5.7	506	4	US-09-107-532A-5994	Sequence 5994, Ap
35	100.5	5.6	478	4	US-09-328-352-6512	Sequence 6512, Ap
36	100.5	5.6	1388	2	US-08-685-576-1	Sequence 1, Appli
37	100	5.6	460	3	US-08-942-012B-33	Sequence 33, Appli
38	100	5.6	831	4	US-09-269-861A-8	Sequence 8, Appli
39	99	5.6	2954	4	US-09-150-867-1	Sequence 1, Appli
40	98.5	5.5	568	4	US-09-710-279-2482	Sequence 2482, Ap
41	98.5	5.5	967	4	US-09-543-681A-6407	Sequence 6407, Ap
42	98.5	5.5	1145	4	US-09-710-279-1432	Sequence 1432, Ap
43	98.5	5.5	1151	3	US-09-134-001C-3242	Sequence 3242, Ap
44	98	5.5	1388	4	US-09-976-594-296	Sequence 296, App
45	98	5.5	2662	4	US-09-595-684B-31	Sequence 31, Appli

ALIGNMENTS

RESULT 1
US-09-190-965-3
; Sequence 3, Application US/09190965
; Patent No. 6071721
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PR-0635 US
; CURRENT APPLICATION NUMBER: US/09/190,965
; CURRENT FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE: -
; OTHER INFORMATION: g262934
US-09-190-965-3

Query Match	90.7%	Score	1614.5	DB 3	Length	341			
Best Local Similarity	92.9%	Pred. No.	3.1e-151						
Matches	325	Conservative	2	Mismatches	14	Indels	9	Gaps	2
Qy	1	MPFPFGKSHKSPADIVKNL	ESMAVLEKQDISDKA	KATEEVS	KNLVAMKEILYGTNEK	60			
Db	1	MPFPFGKSHKSPADIVKNL	ESMAVLEKQDISDKA	KATEEVS	KNLVAMKEILYGTNEK	60			
Qy	61	EPQTEVAQAQLAQLYNSGLL	STLVADQLIDFEGKDV	QAQIFNNILRRQIGTRTPTEYI	120				
Db	61	EPQTEVAQAQLAQLYNSGLL	STLVADQLIDFEGKDV	QAQIFNNILRRQIGTRTPTEYI	120				
Qy	121	CTQONILFMLLKGVESPEI	ALNCIMRECIRHEPLAKI	ILWSQFYDFRYVEMSTFDI	180				
Db	121	CTQONILFMLLKGVESPEI	ALNCIMRECIRHEPLAKI	ILWSQFYDFRYVEMSTFDI	180				
Qy	181	ASDAFAFKDILLTRHKLLS	AEFLQHQYDFRFESEKLLH	SENYVTKPSLKLGLLELLDR	240				
Db	181	ASDAFAFKDILLTRHKLLS	AEFLQHQYDFRFESEKLLH	SENYVTKPSLKLGLLELLDR	240				
Qy	241	HNFTIMTKYISKPENLKLM	NLLRDKSRNTQFEAFHV	KVFVFNPNKTQPTLIDILLNKQA	300				
Db	241	HNFTIMTKYISKPENLKLM	NLLRDKSRNTQFEAFHV	KVFVFNPNKTQPTLIDILLNKQT	300				
Qy	301	KLIEFLSKFQNDRTDCMSS	SVPTTNSRVDLRVPRTR	GIRDLKRPAAQEA	350				
Db	301	KLIEFLSKFQNDRTDCMSS	SVPTTNSRVDLRVPRTR	GIRDLKRPAAQEA	341				

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 3734805
US-09-190-965-1

Query Match      72.9%; Score 1297.5; DB 3; Length 337;
Best Local Similarity 76.0%; Pred. No. 6.6e-120;
Matches 263; Conservative 32; Mismatches 38; Indels 13; Gaps 3;

Qy 1 MPPFFGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
Db 4 MPL-FSKSHKNPAEIVKILKDNLALEKQ---DKKTDKASEEVSLSQAMKEILCGTNEK 59
Qy 61 EPQTEAVAQLAQELYNGLLSTLVADLQIDFEGKKDVAQIFNNILRQIGTRTPVEYI 120
Db 60 EPPTTEAVAQLAQELYSGLLVTLIADLQIDFEGKKDVTQIFNNILRQIGTRSPVEYI 119
Qy 121 CTQONILFMLKGYESPEIALNCGIMLRCEIRHEPLAKIILWSEQYDFFRYVEMSTFDI 180
Db 120 SAHPHILFMLKGYEAPQIALRCGIMLRCEIRHEPLAKIILFSNQFRDFPKYVELSTFDI 179
Qy 181 ASDAFATFKDLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKRQSLKLLGELLDR 240
Db 180 ASDAFATFKDLLTRHKLVADFLQYDITFEYKLLQSENIVTKRQSLKLLGELLDR 239
Qy 241 HNFNTIMTKYISKPENLKLMMNLLRDKSRNIQPEAFHVFKVVFVANPNKTOPILDILLKNOA 300
Db 240 HNFNTIMTKYISKPENLKLMMNLLRDKSPNIQPEAFHVFKVVFVANPHKTOPIVEILLKNOP 299
Qy 301 KLIEFLSKFQNDRTDCMSSSVPTNSRVDLRVKPRTRGIRDLPKPA 346
Db 300 KLIEFLSKFQNDRTDCMSSSVPTNSRVDLRVKPRTRGIRDLPKPA 346

RESULT 4
US-09-470-253-1
; Sequence 1, Application US/09470253
; Patent No. 6365371
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/470,253
; CURRENT FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 3734805
US-09-470-253-1

Query Match      72.9%; Score 1297.5; DB 3; Length 337;
Best Local Similarity 76.0%; Pred. No. 6.6e-120;
Matches 263; Conservative 32; Mismatches 38; Indels 13; Gaps 3;

Qy 1 MPPFFGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
Db 4 MPL-FSKSHKNPAEIVKILKDNLALEKQ---DKKTDKASEEVSLSQAMKEILCGTNEK 59
Qy 61 EPQTEAVAQLAQELYNGLLSTLVADLQIDFEGKKDVAQIFNNILRQIGTRTPVEYI 120
Db 60 EPPTTEAVAQLAQELYSGLLVTLIADLQIDFEGKKDVTQIFNNILRQIGTRSPVEYI 119
Qy 121 CTQONILFMLKGYESPEIALNCGIMLRCEIRHEPLAKIILWSEQYDFFRYVEMSTFDI 180
```

```
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE: -
; OTHER INFORMATION: g262934
US-09-470-253-3

Query Match      90.7%; Score 1614.5; DB 3; Length 341;
Best Local Similarity 92.9%; Pred. No. 3.1e-151;
Matches 325; Conservative 2; Mismatches 14; Indels 9; Gaps 2;

Qy 1 MPPFFGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
Db 1 MPPFFGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
Qy 61 EPQTEAVAQLAQELYNGLLSTLVADLQIDFEGKKDVAQIFNNILRQIGTRTPVEYI 120
Db 61 EPQTEAVAQLAQELYNGLLSTLVADLQIDFEGKKDVAQIFNNILRQIGTRTPVEYI 120
Qy 121 CTQONILFMLKGYESPEIALNCGIMLRCEIRHEPLAKIILWSEQYDFFRYVEMSTFDI 180
Db 121 CTQONILFMLKGYESPEIALNCGIMLRCEIRHEPLAKIILWSEQYDFFRYVEMSTFDI 180
Qy 181 ASDAFATFKDLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKRQSLKLLGELLDR 240
Db 181 ASDAFATFKDLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKRQSLKLLGELLDR 240
Qy 241 HNFNTIMTKYISKPENLKLMMNLLRDKSRNIQPEAFHVFKVVFVANPNKTOPILDILLKNOA 300
Db 241 HNFNTIMTKYISKPENLKLMMNLLRDKSRNIQPEAFHVFKVVFVANPNKTOPILDILLKNOT 300
Qy 301 KLIEFLSKFQNDRTDCMSSSVPTNSRVDLRVKPRTRGIRDLPKPAQOE 350
Db 301 KLIEFLSKFQNDRTDCMSSSVPTNSRVDLRVKPRTRGIRDLPKPAQOE 341

RESULT 3
US-09-190-965-1
; Sequence 1, Application US/09190965
; Patent No. 6071721
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/190,965
; CURRENT FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 337
```

```
Db 120 SAHPHLEMLKGVGEAPQIALRCGIMLRECIHREPLAKIILFSNQFRDFFKYVELSTEDI 179
Qy 181 ASDAFATPKDILLTRHKLSAEFLQHYDRFFSEYKLLHSENYVKROSLLKLGELLDR 240
Db 180 ASDAFATPKDILLTRHKLVAFLEQNYDTIFEDYKLLQSENYVKROSLLKLGELLDR 239
Qy 241 INFNTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKQQA 300
Db 240 INFNTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHVKFVFNPNKTPQIVELLKNQP 299
Qy 301 KLIEFLSKFONDRTDCMSSVPTTNSRVDLRVKPTRGIRDLKRP 346
Db 300 KLIEFLSKFQKERTD-----DEQFADEKNYLIKQIRDLKRTA 336
```

```
RESULT 5
US-09-190-965-4
; Sequence 4, Application US/09190965
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/190,965
; CURRENT FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE: -
; OTHER INFORMATION: g1794137
US-09-190-965-4
```

```
Query Match 62.7%; Score 1117; DB 3; Length 339;
Best Local Similarity 69.0%; Pred. No. 4.7e-102;
Matches 220; Conservative 51; Mismatches 40; Indels 8; Gaps 5;

Qy 1 MPFPFGSKSHKSPADIVKNLKSMAVLEKQDISDKKAERKATBEVSKNLVAMKEIILYGTNEK 60
Db 1 MPL-FGKSQKSPVELVKSLEAINALE--AGDRKVEKAQEDVSKNLVSIKNMLHGSSDA 56
Qy 61 EPQTE-AVAQLAQELYNLSGLLSTLVADLQIDFEGKDVQAOIFNNILRQIGTRPTVEY 119
Db 57 EPPADYVVAQSQELYNLSGLLSTLVADLQIDFEGKDVQAOIFNNILRQIGTRPTVEY 116
```

```
Qy 120 ICTQONILFMLLKGVES--PEIALNCGIMLRECIHREPLAKIILWSEQYDFFRVEMST 177
Db 117 ICTPEILFTLMAGYEDAHPEIALNSGTMLECARYEALAKIMLHSDDEFFKFFRVEVST 176
Qy 178 FDIASDAFATPKDILLTRHKLSAEFLQHYDRFFSE-YEKLLHSENYVKROSLLKLGEL 236
Db 177 FDIASDAFATPKDILLTRHKLSAEFLDANYDKFFSQHYQRLNSENYYVTRQSLKLGEL 236
Qy 237 LLDRHNTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILL 296
Db 237 LLDRHNTVMTYISSEPNLKLMMNMLKESRNIQFEAFHVKFVFNPNKPKPILDILL 296
Qy 297 KNOAKLIEFLSKFQNDRTD 315
Db 297 RNQTKLVDFLTNFTDRSE 315
```

```
RESULT 6
US-09-470-253-4
; Sequence 4, Application US/09470253
; Patent No. 6365371
; GENERAL INFORMATION:
```

```
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/470,253
; CURRENT FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE: -
; OTHER INFORMATION: g1794137
US-09-470-253-4
```

```
Query Match 62.7%; Score 1117; DB 3; Length 339;
Best Local Similarity 69.0%; Pred. No. 4.7e-102;
Matches 220; Conservative 51; Mismatches 40; Indels 8; Gaps 5;

Qy 1 MPFPFGSKSHKSPADIVKNLKSMAVLEKQDISDKKAERKATBEVSKNLVAMKEIILYGTNEK 60
Db 1 MPL-FGKSQKSPVELVKSLEAINALE--AGDRKVEKAQEDVSKNLVSIKNMLHGSSDA 56
Qy 61 EPQTE-AVAQLAQELYNLSGLLSTLVADLQIDFEGKDVQAOIFNNILRQIGTRPTVEY 119
Db 57 EPPADYVVAQSQELYNLSGLLSTLVADLQIDFEGKDVQAOIFNNILRQIGTRPTVEY 116
Qy 120 ICTQONILFMLLKGVES--PEIALNCGIMLRECIHREPLAKIILWSEQYDFFRVEMST 177
Db 117 ICTPEILFTLMAGYEDAHPEIALNSGTMLECARYEALAKIMLHSDDEFFKFFRVEVST 176
Qy 178 FDIASDAFATPKDILLTRHKLSAEFLQHYDRFFSE-YEKLLHSENYVKROSLLKLGEL 236
Db 177 FDIASDAFATPKDILLTRHKLSAEFLDANYDKFFSQHYQRLNSENYYVTRQSLKLGEL 236
Qy 237 LLDRHNTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILL 296
Db 237 LLDRHNTVMTYISSEPNLKLMMNMLKESRNIQFEAFHVKFVFNPNKPKPILDILL 296
Qy 297 KNOAKLIEFLSKFQNDRTD 315
Db 297 RNQTKLVDFLTNFTDRSE 315
```

```
RESULT 7
US-09-190-965-5
; Sequence 5, Application US/09190965
; Patent No. 6071721
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/190,965
; CURRENT FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE: -
; OTHER INFORMATION: g1255838
US-09-190-965-5
```

Query Match 59.2%; Score 1054.5; DB 3; Length 377;

Best Local Similarity 62.5%; Pred. No. 8.2e-96;
Matches 205; Conservative 49; Mismatches 61; Indels 13; Gaps 2;

QY 1 MPPFGKSHKSPADIVKNLKSMAVLEK-----QDISDKAEKATEEVSKNLVAM 50
DB 1 MLLFGKSHKSPADVVKTRLEVLITLDKLPFPKLDKQGNIQSKDKYDKALDEVSKNVAMI 60

QY 51 KEILYGTNEKEPOTE---AVAQIAQELYNGLSTLIVADLQIDFEGKXKDVQAIFNNILR 107
DB 61 KSFYGNDSAEPSSEHVHVQVLAQAEVYNANILPMLIKMLPKPEFECKXKDVQGFNNLLR 120

QY 108 RQIGTRPTVEYICTQONILFMLLKGYESPEIALNCGIMLRBECIRHEPLAKILLWSEQFY 167
DB 121 RQIGTRSPVEYLGAPEILLIQVQSVDPDIALTCGLMLRESIRHDLAKILLYSDFVY 180

QY 168 DFRYVEMSTEDIASDAFATKDLTRHKLKLSAEFLQHQYDRFFSEYEKLLHSENYVTKR 227
DB 181 TFLYVQSEVEDISSDAFSTFKELTTRHKAILAEFLDSNYDTFFAQONLLNSKNYVTR 240

QY 228 QSLKLGELLDRHNTMTKYISKPENLKMNNLLRDKSRNIOFAFHVFKVFNANPNK 287
DB 241 QSLKLGELLDRHNTMTKYISNPDNLRMLMELLRDKSRNIOYEAFAHVFKVFNANPNK 300

QY 288 TOPILDILLKNOAKLIEFLSKFQNDRTD 315
DB 301 PKPISDILNRNREKLVEFLSEPHNDRTD 328

RESULT 8

US-09-470-253-5
; Sequence 5, Application US/09470253
; Patent No. 6365371
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/470,253
; PRIOR FILING DATE: 1999-12-22
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE: -
; OTHER INFORMATION: g1255838
US-09-470-253-5

Query Match 59.2%; Score 1054.5; DB 3; Length 377;
Best Local Similarity 62.5%; Pred. No. 8.2e-96;
Matches 205; Conservative 49; Mismatches 61; Indels 13; Gaps 2;

QY 1 MPPFGKSHKSPADIVKNLKSMAVLEK-----QDISDKAEKATEEVSKNLVAM 50
DB 1 MLLFGKSHKSPADVVKTRLEVLITLDKLPFPKLDKQGNIQSKDKYDKALDEVSKNVAMI 60

QY 51 KEILYGTNEKEPOTE---AVAQIAQELYNGLSTLIVADLQIDFEGKXKDVQAIFNNILR 107
DB 61 KSFYGNDSAEPSSEHVHVQVLAQAEVYNANILPMLIKMLPKPEFECKXKDVQGFNNLLR 120

QY 108 RQIGTRPTVEYICTQONILFMLLKGYESPEIALNCGIMLRBECIRHEPLAKILLWSEQFY 167
DB 121 RQIGTRSPVEYLGAPEILLIQVQSVDPDIALTCGLMLRESIRHDLAKILLYSDFVY 180

QY 168 DFRYVEMSTEDIASDAFATKDLTRHKLKLSAEFLQHQYDRFFSEYEKLLHSENYVTKR 227
DB 181 TFLYVQSEVEDISSDAFSTFKELTTRHKAILAEFLDSNYDTFFAQONLLNSKNYVTR 240

QY 228 QSLKLGELLDRHNTMTKYISKPENLKMNNLLRDKSRNIOFAFHVFKVFNANPNK 287
DB 241 QSLKLGELLDRHNTMTKYISNPDNLRMLMELLRDKSRNIOYEAFAHVFKVFNANPNK 300

QY 288 TOPILDILLKNOAKLIEFLSKFQNDRTD 315
DB 301 PKPISDILNRNREKLVEFLSEPHNDRTD 328

RESULT 9
US-09-248-796A-14303
; Sequence 14303, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14303
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14303

Query Match 20.1%; Score 358.5; DB 4; Length 165;
Best Local Similarity 48.6%; Pred. No. 1.5e-27;
Matches 72; Conservative 32; Mismatches 43; Indels 1; Gaps 1;

QY 96 KOVAQIFNNILRQIGTRPTVEYIC-TOONILFMLLKGYESPEIALNCGIMLRBECIRHE 154
DB 13 KOVILFLLLRRTMANKSPTVDYLVHSKPEITMLIKGENLEIGLCQILRDCIKFE 72

QY 155 PLAKIILWSEQFYDFRYVEMSTFDIASDAFATKDLTRHKLKLSAEFLQHQYDRFFSEY 214
DB 73 VINKEVLYSPSFYNFYKQVQIPTEDIATDAMTLLHLLTTRHKLKLSAEFLQHQYDRFFSEY 132

QY 215 EKLHSENYVTKRQSLKLGELLDRHN 242
DB 133 NKLVTSKNYVTKRLSVKLLDELVSQRSN 160

RESULT 10

US-09-248-796A-15202
; Sequence 15202, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15202
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15202

Query Match 9.3%; Score 166; DB 4; Length 125;
Best Local Similarity 34.2%; Pred. No. 1.1e-08;
Matches 38; Conservative 28; Mismatches 43; Indels 2; Gaps 1;

Qy	1	MPPFPKSHKSPADIVKVNKLESMAVLEKQDISDKKAEKATEEVSKNLVAMKEILYGTNEK	60
Db	16	MAELFKENKPTPELVRALNDQ--VLKLDYASPDNAKQYQDECARVLKNNKVVILHGDEV	73
Qy	61	EPOTEAVQAQLAQLNSGLLSTLVAIDLQIDFEKKDQVAQIENNILRRQIG	111
Db	74	EPOPDQITQAQELISTYDCLYLVNLRKLDLFDSDRRCRCYIVSDIATNYG	124
RESULT 11			
US-09-248-796A-15201			
; Sequence 15201, Application US/09248796A			
; Patent No. 6747137			
; GENERAL INFORMATION:			
; APPLICANT: Keith Weinstock et al			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN			
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 107196.132			
; CURRENT APPLICATION NUMBER: US/09/248,796A			
; CURRENT FILING DATE: 1999-02-12			
; PRIOR APPLICATION NUMBER: US 60/074,725			
; PRIOR FILING DATE: 1998-02-13			
; PRIOR APPLICATION NUMBER: US 60/096,409			
; PRIOR FILING DATE: 1998-08-13			
; NUMBER OF SEQ ID NOS: 28208			
; SEQ ID NO 15201			
; LENGTH: 90			
; TYPE: PRT			
; ORGANISM: Candida albicans			
US-09-248-796A-15201			
Query Match 9.2%; Score 163; DB 4; Length 90;			
Best Local Similarity 50.8%; Pred. No. 1.3e-08;			
Matches 31; Conservative 8; Mismatches 22; Indels 0; Gaps 0;			
Qy	249	YISKPENLKLMMNLLRDKSRNIOFEAPHVFKVFVANPKNTQPILDILLKNOAKLIEFLSK	308
Db	1	FFDDANNLKLTMLLSLDKNLQLEGPHTLKFFVANPKRSQKVTDLIKKANFIEFFKT	60
Qy	309	F 309	
Db	61	F 61	
RESULT 12			
US-09-248-796A-20275			
; Sequence 20275, Application US/09248796A			
; Patent No. 6747137			
; GENERAL INFORMATION:			
; APPLICANT: Keith Weinstock et al			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN			
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 107196.132			
; CURRENT APPLICATION NUMBER: US/09/248,796A			
; CURRENT FILING DATE: 1999-02-12			
; PRIOR APPLICATION NUMBER: US 60/074,725			
; PRIOR FILING DATE: 1998-02-13			
; PRIOR APPLICATION NUMBER: US 60/096,409			
; PRIOR FILING DATE: 1998-08-13			
; NUMBER OF SEQ ID NOS: 28208			
; SEQ ID NO 20275			
; LENGTH: 630			
; TYPE: PRT			
; ORGANISM: Candida albicans			
US-09-248-796A-20275			
Query Match 6.5%; Score 115.5; DB 4; Length 630;			
Best Local Similarity 20.5%; Pred. No. 0.012;			
Matches 74; Conservative 80; Mismatches 138; Indels 69; Gaps 15;			
Qy	10	KSPADIVK-NLKESMAVLEKQDISDKKAEKATEEVSK---NLVAMKEILYGTNEKEPQTE	65

```

; NUMBER OF SEQ ID NOS: 5674
;
; SEQ ID NO 4053
; LENGTH: 533
;
; TYPE: PR1
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4053

```

Query Match	6.2%;	Score 110.5;	DB 3;	Length 533;
Best Local Similarity	22.5%;	Pred. NO. 0.029;		
Matches	73;	Conservative 57;	Mismatches 108;	Indels 87; Gaps 18;
QY	14	DI VKNLKE\$MAV -----EKQIDSKKA\$KA\$-TEV\$KNL\$VAMKEI LYGTNEK-----	60	
Db	237	DILRQFSDSIDVLITGHQRD\$ATIKNQ\$T\$ITQPSG\$KTKV\$G VIEYTHDKV\$LIKECN	236	
QY	61	-----EPQTEAVAQ\$LAQELYN\$GL\$TL\$VADLQ I-----DFEGKKD---VAQIF	102	
Db	297	LMNVHNS\$T\$FKPNDE\$DIA-LRNQ\$L--EDW DITQ AELPYAMR INNSFEAK\$SPHAFV\$LL	353	
QY	103	NNILRQIGTRPTV\$EYI\$CTQO\$N L\$M\$K G\$V\$ESP ALNC\$G ML\$EC I\$HE\$PLAK ILW	162	
Db	354	NY LLEK\$G-----AD ACT---AL\$D\$SANG\$DEK-----VTMRD INNP-----	391	
QY	163	SEQFYDFERYVEM\$T\$E\$D ---ASD\$A\$T\$F\$K\$D\$LL\$TR\$K\$LL\$GAE\$FUE---QHVD-RFF\$EY\$EK	216	
Db	392	---F\$M\$T\$K\$V E L\$G\$K\$D K A E R\$S\$Y\$D\$F D V\$N\$K I\$TV\$N\$K\$F\$E P\$P\$Q\$H\$N\$D F\$A\$G QY	448	
QY	217	LL\$H\$EN\$Y\$V\$T\$K\$R\$O\$S\$K\$L\$G\$E L\$LL\$DR\$N-F\$T\$M\$T\$K-----Y S\$P\$K-----ENL	256	
Db	449	TIHV\$P\$H\$Y\$G\$E\$R\$V\$D\$LL N\$A\$P\$L\$Q\$D\$Q Y I\$C\$V\$N\$Y\$R\$A\$V\$G\$G\$N\$Y\$M\$V\$N\$K\$P\$V K D Q E\$G\$A	508	
QY	257	KLM\$N\$LL---RDK\$SR\$N Q\$F\$E\$A\$F\$V\$K	279	
Db	509	QLL D Y L\$N\$H\$D\$LSQ P\$Q V D\$F\$N\$V\$K	533	

Search completed: October 22, 2004, 09:13:52
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2004, 08:55:18 ; Search time 157 Seconds
(without alignments)
799.715 Million cell updates/sec

Title: US-10-089-688-2

Perfect score: 1781

Sequence: 1 MPFFGKSHKSPADIVRNKLK.....RVKPRTRGIRDLKRPAAQQA 350

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04.*

1: geneseqp1980s.*

2: geneseqp1990s.*

3: geneseqp2000s.*

4: geneseqp2001s.*

5: geneseqp2002s.*

6: geneseqp2003as.*

7: geneseqp2003bs.*

8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1781	100.0	350	4	AAB20387 Human acu
2	1635.5	91.8	341	4	AAB48970 Human ANI
3	1635.5	91.8	341	7	ADJ69105 Human hea
4	1635.5	91.8	496	4	AAE10858 Gal14-huma
5	1635.5	91.8	552	4	AAE10859 LexA-huma
6	1614.5	90.7	341	3	AAE94248 Mouse cal
7	1597.5	89.7	354	4	ABG23844 Novel hum
8	1297.5	72.9	337	3	AAE94247 Human cal
9	1297.5	72.9	337	4	AAE82090 Human pol
10	1297.5	72.9	337	4	AAE39078 Human pol
11	1153.5	64.8	289	4	AAE94139 Human pro
12	1123	63.1	339	4	ABB60392 Drosophil
13	1117	62.7	339	3	AAE94249 Drosophil
14	1034.5	59.2	377	3	AAE94250 C. elegan
15	888.5	49.9	237	4	AAE40864 Human pol
16	705.5	39.6	343	3	AAE45273 Arabidops
17	674	37.8	326	3	AAE51053 Arabidops
18	674	37.8	400	3	AAE51052 Arabidops
19	674	37.8	504	3	AAE51051 Arabidops
20	673.5	37.8	300	3	AAE23886 Arabidops
21	669.5	37.6	345	3	AAE05089 Arabidops
22	666.5	37.4	300	3	AAE45274 Arabidops
23	666.5	37.4	300	3	AAE30714 Arabidops
24	666.5	37.4	305	3	AAE30713 Arabidops
25	658.5	37.0	290	3	AAE23887 Arabidops

ALIGNMENTS

RESULT 1

AAB20387

ID AAB20387 standard; protein; 350 AA.

XX AC AAB20387;

XX DT 11-JUN-2001 (first entry)

XX DE Human acute neuronal induced calcium binding protein ANIC-BP-1B.

XX KW Acute neuronal induced calcium binding protein; ANIC-BP-1B;

XX KW spore variant; human; stroke; head trauma; Parkinson's disease;

XX KW Alzheimer's disease; multiple sclerosis; spinal cord injury;

XX KW cerebroprotective; antiparkinsonian; neurotropic; neuroprotective; therapy;

XX KW diagnosis; vaccine.

XX OS Homo sapiens.

XX FN WO200125423-A1.

XX PD 12-APR-2001.

XX PF 28-SEP-2000; 2000WO-EP009475.

XX PR 04-OCT-1999; 99EP-00119113.

XX PA (MERE) MERCK PATENT GMBH.

XX XX Duecker K, Den Daas I;

XX PI WPI; 2001-266306/27.

XX DR N-PSDB; AAF30688.

XX XX Novel human acute neuronal induced calcium-binding protein like protein

XX XX splice variant, useful for treating stroke, acute head trauma,

XX XX Parkinson's disease, Alzheimer's disease multiple sclerosis, spinal cord

XX XX injury.

XX PS Claim 2; Page 44-45; 49pp; English.

XX CC The present sequence is that of a novel human acute neuronal induced

XX CC calcium binding protein-like protein splice variant, ANIC-BP-1B. The

XX CC protein shows homology to other members of the calcium binding protein

XX CC family, including ANIC-BP, a protein discovered by mRNA differential

XX CC display that is upregulated in a rat model of head trauma. ANIC-BP and

XX CC ANIC-BP-1B differ in their C-terminal portions. The variant protein could

XX CC serve as a novel drug target. The invention provides ANIC-BP-1B

XX CC polynucleotides (see AAF30688) and polypeptides, expression vectors, host

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

CC cells and antibodies, as well as methods for producing the protein and
CC for treating or preventing disorders associated with expression of the
CC protein by inhibiting or activating the action of ANIC-BP-1B. Diseases
CC that may be treated include stroke and acute head trauma, Parkinson's
CC disease, Alzheimer's disease, multiple sclerosis and spinal cord injury.
CC The polynucleotides and polypeptides can also be used in diagnostic
CC assays and in vaccines, and to identify agonists and antagonists useful
CC for treating conditions associated with ANIC-BP-1B imbalance
XX
SQ Sequence 350 AA;

Query Match 100.0%; Score 1781; DB 4; Length 350;
Best Local Similarity 100.0%; Pred. No. 6.8e-157;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPPFGKSHKSPADIVKNLKESMAVLEKODISDKAEKATEEYSKNLVAMKEILYGTNEK 60
DB 1 MPPFGKSHKSPADIVKNLKESMAVLEKODISDKAEKATEEYSKNLVAMKEILYGTNEK 60

QY 61 EPOTEVAQAQLAQLYNSGLLSTLVADILQIDFEGKDVQVAFNNILRRQIGTRTPVEYI 120
DB 61 EPOTEVAQAQLAQLYNSGLLSTLVADILQIDFEGKDVQVAFNNILRRQIGTRTPVEYI 120

QY 121 CTQONILFMLLKGYESPEIANGCIGMLRECIHREPLAKIILWSEQYDFPRYVEMSTFDI 180
DB 121 CTQONILFMLLKGYESPEIANGCIGMLRECIHREPLAKIILWSEQYDFPRYVEMSTFDI 180

QY 181 ASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYKLLHSENVYTKRQSLKLGELLDDR 240
DB 181 ASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYKLLHSENVYTKRQSLKLGELLDDR 240

QY 241 HNFITMTKYISKPENLKMNNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQA 300
DB 241 HNFITMTKYISKPENLKMNNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQA 300

RESULT 2
AAB48970
ID AAB48970 standard; protein; 341 AA.
XX
AC AAB48970;
XX
DT 27-MAR-2001 (first entry)
XX
DE Human ANIC-BP (acute neuronal induced calcium-binding protein).
XX
KW Human; acute neuronal induced calcium-binding protein; ANIC-BP;
KW Mo25 homologue; Hyma homologue; drug screening; stroke;
KW acute head trauma; multiple sclerosis; spinal cord injury; vaccine;
KW cerebroprotective; neuroprotective.
XX
OS Homo sapiens.
XX
PN WO200078947-A1.
XX
PD 28-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-EP005457.
XX
PR 22-JUN-1999; 99EP-00112024.
XX
XX (MERE) MERCK PATENT GMBH.
XX
PI Den Daas I, Fischer V, Seyfried C, Von Melchner L;
XX
DR WPI; 2001-102721/11.
DR N-PSDB; AAC91772.
XX
PT Novel acute neuronal induced calcium binding protein, useful for treating

PT acute head trauma, stroke, multiple sclerosis and spinal cord injury.
XX
PS Claim 2; Page 37; 50pp; English.
XX
CC The invention relates to human acute neuronal induced calcium-binding
CC protein (ANIC-BP) and to nucleic acid encoding it. The invention also
CC relates to expression systems and recombinant host cells comprising ANIC-
CC BP DNA, the recombinant production of ANIC-BP, antibodies specific for
CC ANIC-BP, fusion proteins comprising ANIC-BP and an immunoglobulin FC
CC region, and methods of screening for modulators of ANIC-BP proteins.
CC -BP has homology and structural similarity to Hyma and Mo25 proteins.
CC ANIC-BP proteins and nucleotides are useful for treating stroke and acute
CC head trauma, multiple sclerosis and spinal cord injury. ANIC-BP proteins
CC are useful in screening assays, for identifying membrane bound or soluble
CC receptors, and also in vaccines. ANIC-BP nucleotides are useful as
CC diagnostic reagents, as tools for tissue expression studies, for
CC chromosome localisation studies, as genetic vaccines, and in the
CC generation of transgenic animals. The present sequence represents human
CC ANIC-BP
XX
SQ Sequence 341 AA;

Query Match 91.8%; Score 1635.5; DB 4; Length 341;
Best Local Similarity 94.0%; Pred. No. 2.3e-143;
Matches 329; Conservative 1; Mismatches 11; Indels 9; Gaps 2;

QY 1 MPPFGKSHKSPADIVKNLKESMAVLEKODISDKAEKATEEYSKNLVAMKEILYGTNEK 60
DB 1 MPPFGKSHKSPADIVKNLKESMAVLEKODISDKAEKATEEYSKNLVAMKEILYGTNEK 60

QY 61 EPOTEVAQAQLAQLYNSGLLSTLVADILQIDFEGKDVQVAFNNILRRQIGTRTPVEYI 120
DB 61 EPOTEVAQAQLAQLYNSGLLSTLVADILQIDFEGKDVQVAFNNILRRQIGTRTPVEYI 120

QY 121 CTQONILFMLLKGYESPEIANGCIGMLRECIHREPLAKIILWSEQYDFPRYVEMSTFDI 180
DB 121 CTQONILFMLLKGYESPEIANGCIGMLRECIHREPLAKIILWSEQYDFPRYVEMSTFDI 180

QY 181 ASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYKLLHSENVYTKRQSLKLGELLDDR 240
DB 181 ASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYKLLHSENVYTKRQSLKLGELLDDR 240

QY 241 HNFITMTKYISKPENLKMNNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQA 300
DB 241 HNFITMTKYISKPENLKMNNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQA 300

QY 301 KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVKPRTRGIRDLPKPAQOEA 350
DB 301 KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVKPRTRGIRDLPKPAQOEA 350

RESULT 3
ADJ69105
ID ADJ69105 standard; protein; 341 AA.
XX
AC ADJ69105;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human heat mitochondrial protein as a therapeutic target SeqID911.
XX
KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis; LHON;
KW Leber's hereditary optic neuropathy; lactic acidosis and stroke; MELAS;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003087768-A2.
XX

PD 23-OCT-2003.
 XX 04-APR-2003; 2003WO-US010870.
 XX 12-APR-2002; 2002US-0372843P.
 XX 17-JUN-2002; 2002US-0389987P.
 XX 20-SEP-2002; 2002US-0412418P.
 XX (MITO-) MITOKOR.
 XX (BUCK-) BUCK INST AGE RES.
 XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
 XX Warnock DE;
 XX WPI; 2003-845369/78.
 XX Identifying a mitochondrial target for drug screening assays and for
 XX treating diseases associated with altered mitochondrial function,
 XX comprises detecting a modified polypeptide in a sample and correlating
 XX with the disease.
 XX Claim 1; SEQ ID NO 911; 180pp; English.
 XX This invention relates to novel mitochondrial targets that can be used
 XX for therapeutic intervention in treating a disease associated with
 XX altered mitochondrial function. Specifically, it refers to a method for
 XX identifying proteins of the human heart mitochondrial proteome that are
 XX useful for drug screening assays, as well as therapeutic targets. The
 XX present invention describes a method for identifying such proteins that
 XX can be used in the treatment of various diseases associated with altered
 XX mitochondrial function including diabetes mellitus, Huntington's disease,
 XX osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 XX encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 XX ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 XX compositions have neuroprotective, neurotropic, antidiabetic,
 XX anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 XX cytosolic activities. This polypeptide sequence is a human heart
 XX mitochondrial protein of the invention.
 XX Sequence 341 AA;
 Query Match 91.8%; Score 1635.5; DB 7; Length 341;
 Best Local Similarity 94.0%; Pred. No. 2.3e-143;
 Matches 329; Conservative 1; Mismatches 11; Indels 9; Gaps 2;
 Qy 1 MPFFPGSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVSKNLVAMKEILYGTNEK 60
 Db 1 MPFFPGSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVSKNLVAMKEILYGTNEK 60
 Qy 61 EPQTEAVAQAELYNGLSTLVADLQIDFEGKQDVAFNNILRRQIGTRTPTVEYI 120
 Db 61 EPQTEAVAQAELYNGLSTLVADLQIDFEGKQDVAFNNILRRQIGTRTPTVEYI 120
 Qy 121 CTQONILFMLKGVESPEIALNCGIMLRCEIRHPEPLAKIILWSEQYDFFRYVEMSTFDI 180
 Db 121 CTQONILFMLKGVESPEIALNCGIMLRCEIRHPEPLAKIILWSEQYDFFRYVEMSTFDI 180
 Qy 181 ASDAFATFKOLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKQSLKLGELLDDR 240
 Db 181 ASDAFATFKOLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKQSLKLGELLDDR 240
 Qy 241 HNFITMTYISKPENLKMNNLLRDKSRNIOFEAFHVKFVFNPNKTPQILDILLKNOA 300
 Db 241 HNFITMTYISKPENLKMNNLLRDKSRNIOFEAFHVKFVFNPNKTPQILDILLKNOA 300
 Qy 301 KLIEFLSKFQNDRTDCWSSVPTTNSRVDLVKPTRGIRDLKPAQOEA 350
 Db 301 KLIEFLSKFQNDRTDCWSSVPTTNSRVDLVKPTRGIRDLKPAQOEA 350
 RESULT 4
 ID AAE10858 standard; protein; 496 AA.

XX AAE10858;
 XX 18-DEC-2001 (first entry)
 XX Gal4-human ANIC-BP-1 fusion protein.
 XX Human; acute neuronal induced calcium binding protein type 1 ligand;
 XX ANIC-BP-1; human disease; stroke; head trauma; multiple sclerosis;
 XX Parkinson's disease; Alzheimer's disease; spinal cord injury; vaccine;
 XX gene therapy; fusion protein; Gal4 protein.
 XX Homo sapiens.
 XX Unidentified.
 XX Chimeric.
 XX WO200170771-A2.
 XX 27-SEP-2001.
 XX 20-MAR-2001; 2001WO-EP003149.
 XX 21-MAR-2000; 2000EP-00106110.
 XX (MERE) MERCK PATENT GMBH.
 XX Den Daas I, Duecker K, Hock B;
 XX WPI; 2001-607519/69.
 XX Novel acute neuronal induced calcium binding protein type 1 ligand
 XX polypeptides, useful in the treatment of stroke, head trauma, multiple
 XX sclerosis, Parkinson's disease, Alzheimer's disease and spinal cord
 XX injury.
 XX Disclosure; Page 42-44; 46pp; English.
 XX The invention relates to human acute neuronal induced calcium binding
 XX protein type 1 (ANIC-BP-1) ligand polypeptides and polynucleotides.
 XX Sequences of the invention are useful for treating human diseases
 XX including stroke, head trauma, multiple sclerosis, Parkinson's disease,
 XX Alzheimer's disease and spinal cord injury. They are also useful as
 XX vaccines. ANIC-BP-1 ligands are useful for identifying membrane bound
 XX soluble receptors. Polynucleotides of the invention are useful as
 XX diagnostic reagents, for chromosome localization studies and as valuable
 XX tools for tissue expression studies. They are also useful in gene
 XX therapy. The present sequence is Gal4-human ANIC-BP-1 fusion protein
 XX comprising the Gal4 protein and a C-terminally linked human ANIC-BP-1
 XX protein
 XX Sequence 496 AA;
 Query Match 91.8%; Score 1635.5; DB 4; Length 496;
 Best Local Similarity 94.0%; Pred. No. 3.9e-143;
 Matches 329; Conservative 1; Mismatches 11; Indels 9; Gaps 2;
 Qy 1 MPFFPGSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVSKNLVAMKEILYGTNEK 60
 Db 156 MPFFPGSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVSKNLVAMKEILYGTNEK 215
 Qy 61 EPQTEAVAQAELYNGLSTLVADLQIDFEGKQDVAFNNILRRQIGTRTPTVEYI 120
 Db 216 EPQTEAVAQAELYNGLSTLVADLQIDFEGKQDVAFNNILRRQIGTRTPTVEYI 275
 Qy 121 CTQONILFMLKGVESPEIALNCGIMLRCEIRHPEPLAKIILWSEQYDFFRYVEMSTFDI 180
 Db 276 CTQONILFMLKGVESPEIALNCGIMLRCEIRHPEPLAKIILWSEQYDFFRYVEMSTFDI 335
 Qy 181 ASDAFATFKOLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKQSLKLGELLDDR 240
 Db 336 ASDAFATFKOLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKQSLKLGELLDDR 395
 Qy 241 HNFITMTYISKPENLKMNNLLRDKSRNIOFEAFHVKFVFNPNKTPQILDILLKNOA 300

Db 396 HNFMTKVISKPENLKMNNLLRDKSRNIQFEAFHFVKFVANPNKTOPILDILLKNOA 455
QY 301 KLIEFLSKFQNDRTDCMSSSVPTNSRVDLRVKPRTRGIRDLKRPAAQOEA 350
Db 456 KLIEFLSKFQNDRT-----DEQFNDEKTYLVKQ-----IRDLKRPAAQOEA 496

RESULT 5
AAE10859
ID AAE10859 standard; protein; 552 AA.

XX AAE10859;
DT 18-DEC-2001 (first entry)
XX LexA-human ANIC-BP-1 fusion protein.
XX Human; acute neuronal induced calcium binding protein type 1 ligand;
KW ANIC-BP-1; human disease; stroke; head trauma; multiple sclerosis;
KW Parkinson's disease; Alzheimer's disease; spinal cord injury; vaccine;
KW gene therapy; fusion protein; LexA protein.

XX Homo sapiens.
OS Unidentified.
OS Chimeric.

PH Key Location/Qualifiers
FT Region i. .202
FT /note= "LexA protein"
FT 203. .552
FT /note= "Human ANIC-BP-1 protein"

XX WO200170771-A2.
XX 27-SEP-2001.
XX 20-MAR-2001; 2001WO-EP003149.
XX 21-MAR-2000; 2000EP-00106110.
XX (MERE) MERCK PATENT GMBH.

XX Den Daas I, Duecker K, Hock B;
XX WPI; 2001-607519/69.

XX Novel acute neuronal induced calcium binding protein type 1 ligand
PT polypeptides, useful in the treatment of stroke, head trauma, multiple
PT sclerosis, Parkinson's disease, Alzheimer's disease and spinal cord
PT injury.

PS Disclosure; Page 44-46; 46pp; English.

XX The invention relates to human acute neuronal induced calcium binding
CC protein type 1 (ANIC-BP-1) ligand polypeptides and polynucleotides.
CC Sequences of the invention are useful for treating human diseases
CC including stroke, head trauma, multiple sclerosis, Parkinson's disease,
CC Alzheimer's disease and spinal cord injury. They are also useful as
CC vaccines. ANIC-BP-1 ligands are useful for identifying membrane bound
CC soluble receptors. Polynucleotides of the invention are useful as
CC diagnostic reagents, for chromosome localization studies, and as valuable
CC tools for tissue expression studies. They are also useful in gene
CC therapy. The present sequence is LexA-human ANIC-BP-1 fusion protein
CC comprising the LexA protein and a C-terminally linked human ANIC-BP-1
CC protein

XX Sequence 552 AA;

Query Match 91.8%; Score 1635.5; DB 4; Length 552;
Best Local Similarity 94.0%; Pred. No. 4.5e-143;
Matches 329; Conservative 1; Mismatches 11; Indels 9; Gaps 2;

QY 1 MPFPFGKSHKSPADIVKNIKESMAVLEKODISDKAEKATEEVSKNLVAMKEILYGTNEK 60
Db 212 MPFPFGKSHKSPADIVKNIKESMAVLEKODISDKAEKATEEVSKNLVAMKEILYGTNEK 271
QY 61 EPQTEAVAQLAQELNSGLLSTLVADQLIDFEGKDVQAIFNNILRRQIGTRTPTVEYI 120
Db 272 EPQTEAVAQLAQELNSGLLSTLVADQLIDFEGKDVQAIFNNILRRQIGTRTPTVEYI 331
QY 121 CTQQNILFMLKGYESPEIALNCGIMLRRCIRHPLAKIILWSEQYDFPRYVEMSTFDI 180
Db 332 CTQQNILFMLKGYESPEIALNCGIMLRRCIRHPLAKIILWSEQYDFPRYVEMSTFDI 391
QY 181 ASDAFATFKDILLTRHKLKLSAEFLQHYDRFFSEYEKLLHSENVVTKRQSLKLLGELLDDR 240
Db 392 ASDAFATFKDILLTRHKLKLSAEFLQHYDRFFSEYEKLLHSENVVTKRQSLKLLGELLDDR 451
QY 241 HNFMTKVISKPENLKMNNLLRDKSRNIQFEAFHFVKFVANPNKTOPILDILLKNOA 300
Db 452 HNFMTKVISKPENLKMNNLLRDKSRNIQFEAFHFVKFVANPNKTOPILDILLKNOA 511
QY 301 KLIEFLSKFQNDRTDCMSSSVPTNSRVDLRVKPRTRGIRDLKRPAAQOEA 350
Db 512 KLIEFLSKFQNDRT-----DEQFNDEKTYLVKQ-----IRDLKRPAAQOEA 552

RESULT 6

AAAY94248
ID AAAY94248 standard; protein; 341 AA.

AC AAAY94248;

DT 10-AUG-2000 (first entry)

XX Mouse calcium binding protein MO25.

XX Mouse; calcium binding protein; cancer; inflammation; MO25; CBP;
KW reproductive disorder; autoimmune disorder; developmental disorder;
KW seizure disorder; immune disorder; infection.

XX Mus sp.

XX WO200029580-A1.

XX 25-MAY-2000.

XX 12-NOV-1999; 99WO-US027027.

XX 13-NOV-1998; 98US-00190965.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Guegler KJ, Corley NC, Gorgone GA;

XX WPI; 2000-387793/33.

XX Human hCBP protein, and the nucleic acid encoding it, useful for e.g.
PT diagnosis, prevention and treatment of cancers, immune, developmental or
PT reproductive disorders.

PS Disclosure; Page 66-67; 72pp; English.

XX The present sequence is the mouse calcium binding protein MO25. It was
CC used in a sequence alignment to identify human calcium binding protein
CC hCBP. The hCBP protein and the gene encoding it are useful for the
CC diagnosis and treatment of the following types of disorder: cancers (such
CC as adenocarcinomas), reproductive disorders (such as infertility,
CC ovulatory defects, endometriosis, disruptions of the oestrus and
CC menstrual cycles, polycystic ovary syndrome and ovarian
CC hyperstimulation), autoimmune disorders (such as benign prostatic
CC hyperplasia and prostatitis), developmental disorders (such as Cushing's
CC syndrome, muscular dystrophy and gonadal dysgenesis), hereditary
CC neuropathies, seizure disorders, immune disorders (such as AIDS,
CC allergies, anaemia, asthma, atherosclerosis, cholecystitis, Crohn's

CC disease, diabetes, Graves' disease, multiple sclerosis, psoriasis,
 CC rheumatoid arthritis, scleroderma, Sjogren's syndrome and ulcerative
 CC colitis), and viral, bacterial, fungal, parasitic, protozoal and
 CC helminthic infections
 XX
 SQ Sequence 341 AA;
 Query Match 90.7%; Score 1614.5; DB 3; Length 341;
 Best Local Similarity 92.9%; Pred. No. 28-141;
 Matches 325; Conservative 2; Mismatches 14; Indels 9; Gaps 2;
 QY 1 MPFFPGKSHKSPADIVKNLKESMAVLEKQDISDKKAETEVSKNLVAMKEILYGTNEK 60
 DB 1 MPFFPGKSHKSPADIVKNLKESMAVLEKQDISDKKAETEVSKNLVAMKEILYGTNEK 60
 QY 61 EPQTEAVAQLAQELNSGLLSTLVADLQIDFEGKQDVAFNNILRQIGTRTPTVEYI 120
 DB 61 EPQTEAVAQLAQELNSGLLSTLVADLQIDFEGKQDVAFNNILRQIGTRTPTVEYI 120
 QY 121 CTQONILFMLKGYSPFIALNCGIMLRCEIRHEPLAKIILWSEQYDFFRYVENSTEDI 180
 DB 121 CTQONILFMLKGYSPFIALNCGIMLRCEIRHEPLAKIILWSEQYDFFRYVENSTEDI 180
 QY 181 ASDAFATFKOLLTRHKLKSAEFLQHYDRFFSEYEKLLHSENYVTKRSLKLGELLDDR 240
 DB 181 ASDAFATFKOLLTRHKLKSAEFLQHYDRFFSEYEKLLHSENYVTKRSLKLGELLDDR 240
 QY 241 HNFNTIMTKYISKPENLKLMMNLLRDKSRNIQPEAFHVKFVFNPNKTPQILDILLKNOA 300
 DB 241 HNFNTIMTKYISKPENLKLMMNLLRDKSRNIQPEAFHVKFVFNPNKTPQILDILLKNOA 300
 QY 301 KLIEFLSKFQNDRTDCMSSVPTTNSRVDLRVKPRTRGIRDLKPAQOEA 350
 DB 301 KLIEFLSKFQNDRTDCMSSVPTTNSRVDLRVKPRTRGIRDLKPAQOEA 350
 RESULT 7
 ABG23844
 ID ABG23844 standard; protein; 354 AA.
 XX -AC ABG23844;
 XX 18-FEB-2002 (first entry)
 XX Novel human diagnostic protein #23835.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 XX 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX N-PSDB; AAS88031.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 XX biodiversity.
 XX
 PS Claim 20; SEQ ID NO 54203; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 354 AA;
 Query Match 89.7%; Score 1597.5; DB 4; Length 354;
 Best Local Similarity 92.0%; Pred. No. 8.2e-140;
 Matches 322; Conservative 2; Mismatches 17; Indels 9; Gaps 2;
 QY 1 MPFFPGKSHKSPADIVKNLKESMAVLEKQDISDKKAETEVSKNLVAMKEILYGTNEK 60
 DB 14 MPFFPGKSHKSPADIVKNLKESMAVLEKQDISDKKAETEVSKNLVAMKEILYGTNEK 73
 QY 61 EPQTEAVAQLAQELNSGLLSTLVADLQIDFEGKQDVAFNNILRQIGTRTPTVEYI 120
 DB 74 DPQTEAGAQAQELNSGLLITLVADLQIDFEGKQDVAFNNILRQIGTRTPTVEYI 133
 QY 121 CTQONILFMLKGYSPFIALNCGIMLRCEIRHEPLAKIILWSEQYDFFRYVENSTEDI 180
 DB 134 CTQONILFMLKGYSPFIALNCGIMLRCEIRHEPLAKIILWSEQYDFFRYVENSTEDI 193
 QY 181 ASDAFATFKOLLTRHKLKSAEFLQHYDRFFSEYEKLLHSENYVTKRSLKLGELLDDR 240
 DB 194 ASDAFATFKOLLTRHKLKSAEFLQHYDRFFSEYEKLLHSENYVTKRSLKLGELLDDR 253
 QY 241 HNFNTIMTKYISKPENLKLMMNLLRDKSRNIQPEAFHVKFVFNPNKTPQILDILLKNOA 300
 DB 254 HNFNTIMTKYISKPENLKLMMNLLRDKSRNIQPEAFHVKFVFNPNKTPQILDILLKNOA 313
 QY 301 KLIEFLSKFQNDRTDCMSSVPTTNSRVDLRVKPRTRGIRDLKPAQOEA 350
 DB 314 KLIEFLSKFQNDRTDCMSSVPTTNSRVDLRVKPRTRGIRDLKPAQOEA 354
 RESULT 8
 AAY94247
 ID AAY94247 standard; protein; 337 AA.
 XX AC AAY94247;
 XX 10-AUG-2000 (first entry)
 XX Human calcium binding protein hCBP.
 XX Human; calcium binding protein; cancer; inflammation; CBP;
 KW reproductive disorder; autoimmune disorder; developmental disorder;
 KW seizure disorder; immune disorder; infection.
 OS Homo sapiens.
 XX WO200029580-A1.
 XX 25-MAY-2000.

Fri Oct 22 11:22:09 2004

us-10-089-688-2.rag

XX 12-NOV-1999; 99WO-US027027.
PF 13-NOV-1998; 98US-00190965.
FR (INCY-) INCYTE PHARM INC.
XX Tang YT, Guegler KJ, Corley NC, Gorgone GA;
XX WPI; 2000-387793/33.
DR N-PSDB; AAA27332.
XX Human hCBP protein, and the nucleic acid encoding it, useful for e.g.
PT diagnosis, prevention and treatment of cancers, immune, developmental or
PT reproductive disorders.
XX
PS Claim 1; Fig 1; 72pp; English.
XX The present sequence is the human calcium binding protein hCBP. It was
CC obtained by screening a coronary artery smooth muscle cDNA library, from
CC which five overlapping nucleic acids were isolated, sequenced and
CC expressed to give the protein. The protein and the gene encoding it are
CC useful for the diagnosis and treatment of the following types of
CC disorder: cancers (such as adenocarcinomas), reproductive disorders (such
CC as infertility, ovulatory defects, polycystic ovary syndrome and ovarian
CC oestrus and menstrual cycles, polycystic ovary syndrome and ovarian
CC hyperstimulation), autoimmune disorders (such as benign prostatic
CC hyperplasia and prostatitis), developmental disorders (such as Cushing's
CC syndrome, muscular dystrophy and gonadal dysgenesis), hereditary
CC neuropathies, seizure disorders, immune disorders (such as AIDS,
CC allergies, anaemia, asthma, atherosclerosis, cholecystitis, Crohn's
CC disease, diabetes, Graves' disease, multiple sclerosis, psoriasis,
CC rheumatoid arthritis, scleroderma, Sjogren's syndrome and ulcerative
CC colitis), and viral, bacterial, fungal, parasitic, protozoal and
CC helminthic infections
XX
SQ Sequence 337 AA;

Query Match 72.9%; Score 1297.5; DB 3; Length 337;
Best Local Similarity 76.0%; Pred. No. 6.4e-112;
Matches 263; Conservative 32; Mismatches 38; Indels 13; Gaps 3;
QY 1 MPFPFGKSHKSPADIVKNLKSMAVLEKQDISDKKAEKATEEYSKNLVAMKEILYGTNEK 60
DB 4 MPL-FSKSHKNPAEIVKILNDLAILLEKQ--DKKTDKASEEVSLSQAMKEILCGTNEK 59
QY 61 EPOTEVAQAQLAEYLYSSGLLSTLVADLQIDFEGKDVQAFNNILRRQIGTRPTVEYI 120
DB 60 EPTEVAQAQAELYSGLLVTIADLQIDFEGKDVQAFNNILRRQIGTRSPTEYI 119
QY 121 CTQONILFMLKGVESPEIANGCIGMLRECIRHEPLAKIILWSQFYDFRYVEMSTFDI 180
DB 120 SAHPHILFMLKGVAEAPQIALRCGIMLRRECIRHEPLAKIILFSNQFRDFKYLVELSTFDI 179
QY 181 ASDAFATFKDILTRHKLLSAEFLRQHYDRFFSEYEKLLHSENYVTKQSLKLLGELLDR 240
DB 180 ASDAFATFKDILTRHKLVADFLQNYDTTFEDYEKLLQSENYVTKQSLKLLGELLDR 239
QY 241 HNFNTIMTKYISKPENLKMNNLLRDKSRNTQFEAFHVKFVFNPNKTOPILDTLLKNQA 300
DB 240 HNFNTIMTKYISKPENLKMNNLLRDKSPNTQFEAFHVKFVFNPNKTOPILDTLLKNQ 299
QY 301 KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVKPTRTGIRDLKRP 346
DB 300 KLIEFLSKFQKERTD-----DEQFADEKNYLIKQIRDLKRTA 336
RESULT 9
AAB82090
ID AAB82090 standard; protein; 337 AA.
XX
AC AAB82090;
XX

DT 26-JUN-2001 (first entry)
XX Human Acute Neuronal Induced Calcium Binding Protein, ANIC-BP.
DE Human; cerebroprotective; neuroprotective; vulnary; vaccine;
XX gene therapy; Acute Neuronal Induced Calcium Binding Protein; ANIC-BP;
KW stroke; acute head trauma; multiple sclerosis; spinal cord injury.
KW
XX Homo sapiens.
XX WO200123552-A1.
XX 05-APR-2001.
PD 18-SEP-2000; 2000WO-EP009132.
PF 24-SEP-1999; 99EP-00118848.
PR (MERE) MERCK PATENT GMBH.
XX Den Daas I, Duecker K;
PI WPI; 2001-308142/32.
XX N-PSDB; AAF86462.
DR Novel human acute neuronal induced calcium binding polypeptide, and
XX polynucleotides encoding them useful for diagnosing or treating stroke,
PT acute head trauma, multiple sclerosis and spinal cord injury.
PT
XX Claim 1; Page 41-42; 45pp; English.
XX The present sequence is the protein sequence for human Acute Neuronal
CC Induced Calcium Binding Protein (ANIC-BP). ANIC-BP coding sequence and
CC protein are useful for treating stroke, acute head trauma, multiple
CC sclerosis and spinal cord injury. ANIC-BP coding sequence and protein are
CC also useful as vaccines for inducing an immunological response in a
CC mammal
XX
SQ Sequence 337 AA;
Query Match 72.9%; Score 1297.5; DB 4; Length 337;
Best Local Similarity 76.0%; Pred. No. 6.4e-112;
Matches 263; Conservative 32; Mismatches 38; Indels 13; Gaps 3;
QY 1 MPFPFGKSHKSPADIVKNLKSMAVLEKQDISDKKAEKATEEYSKNLVAMKEILYGTNEK 60
DB 4 MPL-FSKSHKNPAEIVKILNDLAILLEKQ--DKKTDKASEEVSLSQAMKEILCGTNEK 59
QY 61 EPOTEVAQAQLAEYLYSSGLLSTLVADLQIDFEGKDVQAFNNILRRQIGTRPTVEYI 120
DB 60 EPTEVAQAQAELYSGLLVTIADLQIDFEGKDVQAFNNILRRQIGTRSPTEYI 119
QY 121 CTQONILFMLKGVESPEIANGCIGMLRECIRHEPLAKIILWSQFYDFRYVEMSTFDI 180
DB 120 SAHPHILFMLKGVAEAPQIALRCGIMLRRECIRHEPLAKIILFSNQFRDFKYLVELSTFDI 179
QY 181 ASDAFATFKDILTRHKLLSAEFLRQHYDRFFSEYEKLLHSENYVTKQSLKLLGELLDR 240
DB 180 ASDAFATFKDILTRHKLVADFLQNYDTTFEDYEKLLQSENYVTKQSLKLLGELLDR 239
QY 241 HNFNTIMTKYISKPENLKMNNLLRDKSRNTQFEAFHVKFVFNPNKTOPILDTLLKNQA 300
DB 240 HNFNTIMTKYISKPENLKMNNLLRDKSPNTQFEAFHVKFVFNPNKTOPILDTLLKNQ 299
QY 301 KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVKPTRTGIRDLKRP 346
DB 300 KLIEFLSKFQKERTD-----DEQFADEKNYLIKQIRDLKRTA 336
RESULT 10
AAM39078
ID AAM39078 standard; protein; 337 AA.
XX

CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
XX Sequence 289 AA;
SQ
Query Match 64.8%; Score 1153.5; DB 4; Length 289;
Best Local Similarity 77.8%; Pred. No. 1.3e-98;
Matches 231; Conservative 23; Mismatches 34; Indels 9; Gaps 1;
QY 50 MKELVGTNEKEPQTAVQAQLAQELYNGLSTLVADLQIDFEGKQVQAIFNNILRRQ 109
Db 1 MKELVGTNEKEPQTAVQAQLAQELYNGLSTLVADLQIDFEGKQVQAIFNNILRRQ 60
QY 110 IGRPTTVEICTQNNILFLLKGYSPETALNCGIMLRRCIRHEPLAKIILWSEQFYDF 169
Db 61 IGRSTPVEIISHPHILFLLKGYEAPQALRCGIMLRRCIRHEPLAKIILWSEQFYDF 120
QY 170 FRYVEMSTFDIASDAFATFKDILLTRHKLKLSAEFLQHYDFRFFSEYKLLHSENYVTKROS 229
Db 121 FKVELSTFDIASDAFATFKDILLTRHKLKLSAEFLQHYDFRFFSEYKLLHSENYVTKROS 180
QY 230 LKLGELLLDRHNFTIMTKYISKPENIKLMMNLLRDKSNIQFEAFHVKFVFNPNKTO 289
Db 181 LKLGELLLDRHNFTIMTKYISKPENIKLMMNLLRDKSNIQFEAFHVKFVFNPNKTO 240
QY 290 PILDILKNOAKLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVPRTRGIRDLKRP 346
Db 241 PIVEILKNOAKLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVPRTRGIRDLKRP 288
RESULT 12
AB860392
ID AB860392 standard; protein; 339 AA.
XX
XX AB860392;
DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 7968.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-65860/75.
XX
XX N-PSDB; ABL04495.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
XX Disclosure; SEQ ID NO 7968; 21pp + Sequence Listing; English.
XX
XX

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 339 AA;
SQ
Query Match 63.1%; Score 1123; DB 4; Length 339;
Best Local Similarity 69.3%; Pred. No. 1.1e-95;
Matches 221; Conservative 50; Mismatches 40; Indels 8; Gaps 5;
QY 1 MPFPCKSHKSPADIYKNIKESMAVLEKQDISDKAEKATEEYVSKNLVAMKEILYGTNEK 60
Db 1 MPL-FGKSQKSPVELYKSLKEAINALE---AGDRKVEKQEDVSKNLVSKNMLYSSDA 56
QY 61 EPQTE-AVAQLAQELYNGLSTLVADLQIDFEGKQVQAIFNNILRRQIGTRTPTVEY 119
Db 57 EPPADYVVAQLSQELYNGLSTLVADLQIDFEGKQVQAIFNNILRRQIGTRTPTVEY 116
QY 120 ICTQQNILLFLLKGYES--PEIALNCGIMLRRCIRHEPLAKIILWSEQFYDFRIVEMST 177
Db 117 ICTKPEILFTIMAGYEDAHEPEIALNCGIMLRRCIRHEPLAKIILWSEQFYDFRIVEMST 176
QY 178 FDIASDAFATFKDILLTRHKLKLSAEFLQHYDFRFFSE-YEKLLHSENYVTKROSILKLGEL 236
Db 177 FDIASDAFATFKDILLTRHKLKLSAEFLQHYDFRFFSE-YEKLLHSENYVTKROSILKLGEL 236
QY 237 LLDRHNFTIMTKYISKPENIKLMMNLLRDKSNIQFEAFHVKFVFNPNKTOPIILDILL 296
Db 237 LLDRHNFTIMTKYISKPENIKLMMNLLRDKSNIQFEAFHVKFVFNPNKTOPIILDILL 296
QY 297 KNOAKLIEFLSKFQNDRTD 315
Db 297 RNOTKLVDLTFNFTDRSE 315
RESULT 13
AA94249
ID AA94249 standard; protein; 339 AA.
XX
XX AA94249;
XX
XX 10-AUG-2000 (first entry)
XX
XX Drosophila calcium binding protein DMO25.
XX
XX Drosophila; calcium binding protein; cancer; inflammation; DMO25; CBP;
XX reproductive disorder; autoimmune disorder; developmental disorder;
XX seizure disorder; immune disorder; infection.
XX
XX Drosophila melanogaster.
XX
XX WO200029580-A1.
XX
XX 25-MAY-2000.
XX
XX 12-NOV-1999; 99WO-US027027.
XX
XX 13-NOV-1998; 98US-00190965.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Tang YT, Guegler KJ, Corley NC, Gorgone GA;
XX WPI; 2000-387793/33.
XX
XX Human hCBP protein, and the nucleic acid encoding it, useful for e.g.
XX PT

PT diagnosis, prevention and treatment of cancers, immune, developmental or
XX reproductive disorders.
PS Disclosure; Page 67-68; 72pp; English.

XX The present sequence is the Drosophila calcium binding protein DMO25. It
CC was used in a sequence alignment to identify human calcium binding
CC protein hCBP. The hCBP protein and the gene encoding it are useful for
CC the diagnosis and treatment of the following types of disorder: cancers
CC (such as adenocarcinomas), reproductive disorders (such as infertility,
CC ovulatory defects, endometriosis, disruptions of the oestrus and
CC menstrual cycles, polycystic ovary syndrome and ovarian
CC hyperstimulation), autoimmune disorders (such as benign prostatic
CC hyperplasia and prostatitis), developmental disorders (such as Cushing's
CC syndrome, muscular dystrophy and gonadal dysgenesis), hereditary
CC neuropathies, seizure disorders, immune disorders (such as AIDS,
CC allergies, anaemia, asthma, atherosclerosis, cholecystitis, Crohn's
CC disease, diabetes, Graves' disease, multiple sclerosis, psoriasis,
CC rheumatoid arthritis, scleroderma, Sjogren's syndrome and ulcerative
CC colitis), and viral, bacterial, fungal, parasitic, protozoal and
CC helminthic infections
XX Sequence 339 AA;

Query Match 62.7%; Score 1117; DB 3; Length 339;
Best Local Similarity 69.0%; Pred. No. 4.1e-95;
Matches 220; Conservative 51; Mismatches 40; Indels 8; Gaps 5;
Qy 1 MPFFPGSHKSPADIVKNLKSMAVLEKQDISDKAEKATEEVSQNLVAMKEILYGTNEK 60
Db 1 MPL-FGSKQSPVBLVSLKEAINALE---AGDRKVEKAQEDVSKNLVSIKMLHGSSDA 56
Qy 61 EPOTE-AVAQAELYNGLSTLVADQLIDFEGKQDVQAIQFNILRRQIGTRTPTVEY 119
Db 57 EPPADYVVAQSLQELYNGLSTLVADQLIDFEGKQDVQAIQFNILRRQIGTRTPTVEY 116
Qy 120 ICTQONILFMLLKGYES--PEIALNCGIMLRECIHPEPLAKIILWSEQYDFFRYEMST 177
Db 117 ICTKPEILFTLMAGYEDAHPEIALNCGIMLRECIHPEPLAKIILWSEQYDFFRYEMST 176
Qy 178 FDIASDAFATKDLTRHKLKLSAEFLQHYDRFTSE-YEKLLHSENYTKQSLKLGEL 236
Db 177 FDIASDAFSTFKELLTRHKLKLSAEFLQHYDRFTSE-YEKLLHSENYTKQSLKLGEL 236
Qy 237 LLDRHNFTMTKYISKPNKLMNMLLDKGRNIQFEAFHFKVFNPNKQTPDILLL 296
Db 237 LLDRHNFTMTKYISKPNKLMNMLLDKGRNIQFEAFHFKVFNPNKQTPDILLL 296
Qy 297 KNOAKLIEFLSKFQNDRTD 315
Db 297 RNQTKLVDFLTNFTDRSE 315

RESULT 14
AA94250
ID AA94250 standard; protein; 377 AA.
XX
AC AA94250;
XX
DT 10-AUG-2000 (first entry)
XX
DE C. elegans yeast-like calcium binding protein.
XX
KW Calcium binding protein; cancer; inflammation; yeast-like CBP; CBP;
KW reproductive disorder; autoimmune disorder; developmental disorder;
KW seizure disorder; immune disorder; infection.
XX
OS Caenorhabditis elegans.
XX
PN WO200029580-A1.
XX
PD 25-MAY-2000.
XX

PF 12-NOV-1999; 99WO-US027027.
XX
PR 13-NOV-1998; 98US-00190965.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Guegler KJ, Corley NC, Gorgone GA;
XX WPI; 2000-387793/33.
DR
PT Human hCBP protein, and the nucleic acid encoding it, useful for e.g.
PT diagnosis, prevention and treatment of cancers, immune, developmental or
PT reproductive disorders.
PS Disclosure; Page 68-69; 72pp; English.
XX
CC The present sequence is the C. elegans yeast-like CBP. It was used in a
CC sequence alignment to identify human calcium binding protein hCBP. The
CC hCBP protein and the gene encoding it are useful for the diagnosis and
CC treatment of the following types of disorder: cancers (such as
CC adenocarcinomas), reproductive disorders (such as infertility, ovulatory
CC defects, endometriosis, disruptions of the oestrus and menstrual cycles,
CC polycystic ovary syndrome and ovarian hyperstimulation), autoimmune
CC disorders (such as benign prostatic hyperplasia and prostatitis),
CC developmental disorders (such as Cushing's syndrome, muscular dystrophy
CC and gonadal dysgenesis), hereditary neuropathies, seizure disorders,
CC immune disorders (such as AIDS, allergies, anaemia, asthma,
CC atherosclerosis, cholecystitis, Crohn's disease, diabetes, Graves'
CC disease, multiple sclerosis, psoriasis, rheumatoid arthritis, and viral,
CC scleroderma, Sjogren's syndrome and ulcerative colitis), and viral,
CC bacterial, fungal, parasitic, protozoal and helminthic infections
XX Sequence 377 AA;

Query Match 59.2%; Score 1054.5; DB 3; Length 377;
Best Local Similarity 62.5%; Pred. No. 3.1e-89;
Matches 205; Conservative 49; Mismatches 61; Indels 13; Gaps 2;
Qy 1 MPFFPGSHKSPADIVKNLKSMAVLEK-----QDISDKAEKATEEVSQNLVAM 50
Db 1 MPL-FGSHKSPADIVKNLKSMAVLEK-----QDISDKAEKATEEVSQNLVAM 60
Qy 51 KEILYGTNEKEPQTE---AVAQAELYNGLSTLVADQLIDFEGKQDVQAIQFNILR 107
Db 61 KSFYIGNDSAEPSSEHVQVQAQAEVYANILKMLPKFPECKKQVGFNNLLR 120
Qy 108 RQIGTRPTVEYICTQONILFMLLKGYESPEIALNCGIMLRECIHPEPLAKIILWSEQY 167
Db 121 RQIGTRPTVEYLGARPEILTLVQGYSVDPDIALTCGLMLRESIRHDLAKIILYSDVFY 180
Qy 168 DFFRYVEMSTFDIASDAFATKDLTRHKLKLSAEFLQHYDRFTSEYSEKLLHSENYTKR 227
Db 181 TFFLYVQSEVFDISSDAFSTFKELLTRHKLKLSAEFLQHYDRFTSEYSEKLLHSENYTKR 240
Qy 228 QSLKLGELLLDRHNFTMTKYISKPNKLMNMLLDKGRNIQFEAFHFKVFNPNK 287
Db 241 QSLKLGELLLDRHNFTMTKYISKPNKLMNMLLDKGRNIQFEAFHFKVFNPNK 300
Qy 288 TQPIDILLKNOAKLIEFLSKFQNDRTD 315
Db 301 PKPISDILNRRNREKLVFLSEFHNDRTD 328

RESULT 15
AA40864
ID AA40864 standard; protein; 237 AA.
XX
AC AA40864;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 5795.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2004, 09:03:58 ; Search time 39 seconds
(without alignments)
863.484 Million cell updates/sec

Title: US-10-089-688-2

Perfect score: 1781

Sequence: 1 MPFFPGSKSHKSPADIVKNLK.....RVKPRTRGIRDLKRPAAQEA 350

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1614.5	90.7	341	2	hypothetical calci
2	1054.5	59.2	377	2	hypothetical prote
3	1006	56.5	338	2	hypothetical prote
4	766	43.0	329	2	mo25 homolog [impo
5	666.5	37.4	305	2	hypothetical prote
6	614	34.5	348	2	hypothetical prote
7	501	28.1	399	2	hypothetical prote
8	137.5	7.7	339	2	hypothetical prote
9	122	6.9	602	2	hypothetical prote
10	117	6.6	365	2	hypothetical prote
11	115	6.5	2819	2	conserved hypothet
12	113.5	6.4	1939	2	repeat organellar
13	112.5	6.3	677	2	DNA topoisomerase
14	111	6.2	4981	2	hypothetical prote
15	109	6.1	2166	2	hypothetical prote
16	108.5	6.1	719	2	hypothetical prote
17	108	6.1	2401	2	rhodopy protein -
18	107.5	6.0	1050	2	hypothetical prote
19	106.5	6.0	348	2	hypothetical prote
20	106	6.0	824	2	hypothetical prote
21	105	5.9	809	1	anthrax toxin leth
22	104.5	5.9	617	2	hypothetical prote
23	104.5	5.9	1156	2	hypothetical prote
24	104	5.8	299	1	chromosome assembl
25	104	5.8	592	2	probable arsenical
26	104	5.8	4131	2	hypothetical prote
27	103.5	5.8	415	2	kinase-related tra
28	103.5	5.8	1189	2	AtPase ScII, chrom
29	103.5	5.8	1642	2	NMDA receptor-bind

30 103.5 5.8 2473 1 S38040
31 103 5.8 256 2 B81394
32 103 5.8 355 2 C64644
33 103 5.8 799 2 T00331
34 103 5.8 959 2 T00246
35 102.5 5.8 472 2 T19375
36 102.5 5.8 600 2 T09676
37 102.5 5.8 1300 2 T18364
38 102.5 5.8 1666 2 T38393
39 102 5.7 717 2 T27066
40 102 5.7 717 2 T27067
41 102 5.7 878 2 T23537
42 102 5.7 1026 2 T03108
43 101.5 5.7 1170 2 A56157
44 101 5.7 1285 2 B72420
45 101 5.7 1317 2 A54831

ALIGNMENTS

RESULT 1

157997

hypothetical calcium-binding protein - mouse

C:Species: Mus sp. (mouse)

C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 19-May-2000

C:Accession: I57997

R:Miyamoto, H.; Matsushiro, A.; Nozaki, M.

Mol. Reprod. Dev. 34, 1-7, 1993

A>Title: Molecular cloning of a novel mRNA sequence expressed in cleavage stage mouse em

A:Reference number: I57997; MUID:93119656; PMID:8418809

A:Accession: I57997

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-341 <RES>

A:Cross-references: GB:S51858; NID:g262933; PIDN:AAB24801.1; PID:g262934

C:Superfamily: Saccharomyces hypothetical protein YKL189w

C:Keywords: calcium binding

Query Match 90.7%; Score 1614.5; DB 2; Length 341;

Best Local Similarity 92.9%; Pred. No. 7.2e-99;

Matches 325; Conservative 2; Mismatches 14; Indels 9; Gaps 2;

Qy 1 MPFFPGSKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEVSKNLVAMKEILYGTNEK 60
Db 1 MPFFPGSKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEVSKNLVAMKEILYGTNEK 60
Qy 61 EPQTEVAQAQAELYNGLLSTVLADLQIDFEGKQDVAQIFNNILRRQIGTRTPTVEYI 120
Db 61 EPQTEVAQAQAELYNGLLSTVLADLQIDFEGKQDVAQIFNNILRRQIGTRTPTVEYI 120
Qy 121 CTQONILFMLKGVESPEIALNCGIMLRECIHREPLAKIILWSQFYDFFRYVEMSTFDI 180
Db 121 CTQONILFMLKGVESPEIALNCGIMLRECIHREPLAKIILWSQFYDFFRYVEMSTFDI 180
Qy 181 ASDAFATFKDILLTRKLLISAEFLQHYDRFPFSEYKLLHSENYVTKQSLKLGELLDR 240
Db 181 ASDAFATFKDILLTRKLLISAEFLQHYDRFPFSEYKLLHSENYVTKQSLKLGELLDR 240
Qy 241 HNFITMTKYISKPNLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQITDILLKNQA 300
Db 241 HNFITMTKYISKPNLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQITDILLKNQA 300
Qy 301 KLIEFLSKFQNDRTDCMSSVPTTNSRVDLRVKPRTRGIRDLKRPAAQEA 350
Db 301 KLIEFLSKFQNDRTDCMSSVPTTNSRVDLRVKPRTRGIRDLKRPAAQEA 350

RESULT 2

T16651

hypothetical protein R02E12.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

RESULT 5

G71441
hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 18-Aug-2000
C:Accession: G71441
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ana
C.; Chalwatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: G71441
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-305 <BEV>
A:Cross-references: GB:Z97343; NID:G2245073; PID:e327051; PID:G2245086
C:Genetics:
A:Map position: 4COP9-4G3845
C:Superfamily: Saccharomyces hypothetical protein YKL189w

Query Match 37.4%; Score 666.5; DB 2; Length 305;
Best Local Similarity 44.9%; Pred. No. 1.1e-36;
Matches 136; Conservative 63; Mismatches 93; Indels 11; Gaps 2;

QY 42 EVSKNLVAMKILLVTGNEKEPQTEAVAQAEIYNSGLLSTLVADQLIDFEGKQVQAI 101
DB 8 ELSKIRDLKLLYNSAEFPVAECAQTQFFKADTLRLTLSPNLNLEARKDATQV 67
QY 102 FNNILRQIGTRTPTEVICTQONILFMLKGYSPETALNCGIMLRCEIRHEPLAKIIL 161
DB 68 VANLQROQVNSRLTAADYLESNIDMLDFVGFTNDWALHVGTFRCIRHQIVAKYVL 127
QY 162 WSEQYDFRYVEMSTFDIASDAPATFKDLLTRHKLKSAEFLQHYDRFFSEY-EKLHLS 220
DB 128 DSEHVKKFFYYIQLPNFDIAADAAATFKELLTRHKSSTAEEFLINKEDFFADYNSKLES 187
QY 221 ENYTKRQSLKGLLELLDRHNTIMTKYISKPENLKMMLLRDKNISNIQFEAFHVKV 280
DB 188 TNYTRQAIKLLGDLDRNSAVMTKYVSSMDNLRLIMLLRSSXTIQIEAFHVKL 247
QY 281 FVANPNKTOPILDKLKNQAKLIEFLSKFONDRTDCMSSVPTTNSRVDLRVKPRTRGIR 340
DB 248 FVANQKPSDIANILVANRKLRLILLADIKPKED-----ERFDADKAQVVRFA 297
QY 341 DLK 343
DB 298 NLK 300

RESULT 6

B84448
hypothetical protein At2G03410 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: B84448
R:Liu, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
mu, D.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84448
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-348 <STO>
A:Cross-references: UNIPROT:Q9ZQ77; GB:AE002093; NID:G4333758; PIDN:AADI7435.1; GSPDB:GN
C:Genetics:
A:Gene: At2G03410

A:Map position: 2

C:Superfamily: Saccharomyces hypothetical protein YKL189w

Query Match 34.5%; Score 614; DB 2; Length 348;
Best Local Similarity 39.4%; Pred. No. 3.7e-33;
Matches 125; Conservative 74; Mismatches 112; Indels 6; Gaps 3;

QY 5 FGSHKSPADIVKVLKESMAVLEKQ----DISDKAEKATEEVSKNLVAMKEILYGTNEK 60
DB 5 FPNKSRLLPGEIVRQTRDLIALAESSEETDARNKRLGICAEICRNIRDLKSLIYNGCEA 64
QY 61 EPQTEAVAQAEIYNSGLLSTLVADQLIDFEGKQVQAIFFNNILRQIGTRTPTEYI 120
DB 65 EPVPEACILLTQBFPRADTLRLIKSPIKLEARKDATQIVANLQKQVFRFLVASBYL 124
QY 121 CTQONILFMLKGYE-SPEIALNCGIMLRCEIRHEPLAKIILWSEQYDFRYVEMSTFD 179
DB 125 ESNLDVDSLVGEGDHDHDLALHYTGMLKRCVRQVQVAKYILESKNLEKFFDYVOLPFD 184
QY 180 IASDAFATFKDLLTRHKLKSAEFLQHYDRFFSEY-EKLHSENYVTKRQSLKGLLELL 238
DB 185 VATDASKIFRELLTRHKSSTAEEFLAKNVEFVAEYNTKLEKGSYFKRQASKLLGDVLM 244
QY 239 DRHNTIMTKYISKPENLKMMLLRDKNISNIQFEAFHVKVFNANPKTOPILDKLKN 298
DB 245 DRNSGVVVKYVSSLDNLRIMNLLREPTKNIQLEAPHFIFKLFVANENKPEDIVAILVAN 304
QY 299 QAKLIEFLSKFONDRTD 315
DB 305 RTKILRLFADLKPKED 321

RESULT 7

S34681
hypothetical protein YKL189w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: S34681; S33963; S38021; S38026
R:Wiemann, S.; Voss, H.; Schwagaer, C.; Rupp, T.; Stegemann, J.; Grothues,
submitted to the EMBL Data Library, July 1993
A:Description: Sequencing and analysis of 51.5 kilobases on the left arm of chromosome XI
A:Reference number: S34679
A:Accession: S34681
A:Molecule type: DNA
A:Residues: 1-399 <WIE>
A:Cross-references: UNIPROT:P32464; EMBL:X74151; NID:G450365; PIDN:CAA52249.1; PID:G39523
A:Experimental source: strain S288C
R:Cheret, G.; Mattheakis, L.C.; Sor, F.
Yeast 9, 661-667, 1993
A:Title: DNA sequence analysis of the YCN2 region of chromosome XI in Saccharomyces cere
A:Reference number: S33960; MUID:93348778; PMID:8394042
A:Accession: S33963
A:Molecule type: DNA
A:Residues: 1-399 <CHE>
A:Cross-references: GB:X69765; NID:G296985; PIDN:CAA49422.1; PID:G296989
R:Wiemann, S.; Voss, H.; Schwager, C.; Rupp, T.; Grothues, D.; Senses, C.; Stegemann, J.;
submitted to the Protein Sequence Database, March 1994
A:Reference number: S37825
A:Accession: S38021
A:Molecule type: DNA
A:Residues: 1-399 <WIE>
A:Cross-references: EMBL:Z28189; NID:G486334; PIDN:CAA82032.1; PID:G486335; MIPS:YKL189w
A:Experimental source: strain S288C
R:Maia e Silva, A.; Bossier, P.; Villela, C.; Fernandes, L.; Soares, H.; Guerreiro, P.; R
submitted to the Protein Sequence Database, March 1994
A:Reference number: S38024
A:Accession: S38026
A:Molecule type: DNA
A:Residues: 1-399 <MAI>
A:Cross-references: EMBL:Z28189; NID:G486334; PIDN:CAA82032.1; PID:G486335; MIPS:YKL189w
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:HYM1

A;Cross-references: SGD:S0001672
A;Map position: 111
C;Superfamily: Saccharomycetes hypothetical protein YKL189w

Query Match 28.1%; Score 501; DB 2; Length 399;
Best Local Similarity 33.8%; Pred. No. 1.2e-25;
Matches 110; Conservative 68; Mismatches 137; Indels 10; Gaps 3;

QY 1 MPFPFGSHKSPADIVKNLKESMAVLEKQDISDKAEKATEVSKNIVAMKEILYGTNEK 60
DB 12 MATFWKKNPKTPSDYARLIIEQLNKFSSPSLTQDNKRKVOEECTKYLIGTKHFIVGDTDP 71

QY 61 EPOTEVAQAQALYNSGLSTIVADQLIDFGKDVQAQIFNNILRRQIGTRPTVEYI 120
DB 72 HPTPEAIDELYTAMRADVFYELLHFVDFEAREECMLIFSICLGYSKDNKFVTVYDL 131

QY 121 CTQONILFMLLKGYE-----SPETALNCIMLRCEIRHEPLAKIILWSEQYDFPRYV 173
DB 132 VSQPKTISLMRTAEVALQOKGCDIFLTGNNMIIECIKYEQICRILLKDPQLWKFFFEA 191

QY 174 EMSTFDIADAFATKDLLTRH-KLSAEFLQHYD--RPFSEYKILHSENVVTKRQSL 230
DB 192 KLGNFESTESLIQLSAFTAHPLKYSKEFFSNEINIIRIKCINKLMAHGSYVTKRQST 251

QY 231 KLGELLDRHNTIMTKYISKPENIKLMMNLLRDKSRNIOFAFHVKVVFVANPNKTOP 290
DB 252 KLLASLIVRSNNALMIYINSPENLKLIMTLMTDKSKNLQLEAFNVFVKNVANPRKSKP 311

QY 291 ILDILLKNOAKLIEFLSKFQNDRTD 315
DB 312 VFDILVKNRDKLLITYFTKFLGDSQD 336

RESULT 8
T33477
hypothetical protein T27C10.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33477
R;Zhu, H.J.; Graves, T.; Hawkins, M.
submitted to the EMBL Data Library, October 1998
A;Description: The sequence of C. elegans cosmid T27C10.
A;Reference number: 221354
A;Accession: T33477
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-339 <ZHU>
A;Cross-references: UNIPROT:Q9TZM2; EMBL:AF098504; PIDN:AAC67411.1; GSPDB:GN00019; CESP:
A;Experimental source: strain Bristol N2; clone T27C10
C;Genetics:
A;Gene: CESP:T27C10.3
A;Map position: 1
A;Introns: 72/3; 120/3; 233/3; 295/1

Query Match 7.7%; Score 137.5; DB 2; Length 339;
Best Local Similarity 20.0%; Pred. No. 0.071;
Matches 36; Conservative 43; Mismatches 82; Indels 19; Gaps 3;

QY 178 FDIASDAFAFKDILLTRHKLKSAFLRHOYDRFSEYKELHLSNRYVTKRQSLKLGCELL 237
DB 109 FDIQVTFDILQIIFFTTHESANNFINNLPFRMOTLUKLIACSNFFIQAKSKFLNELF 168

QY 238 LDRHNTIMTKYISKPENIKLMMNLLRDKSRNIOFAFHVKVVFVANPNKTOPILDKL 297
DB 169 TAQINYETRSLWMAEPAFIKLVLAIQSNKHAVSRASVILEIFIRPNRSPEVHEFGR 228

QY 298 NQAKLIEFLSKFQNDRTDCMSSVSP-----TTNSRVDLRVKPRTGIR--DLKRPQOE 349
DB 229 NRNVLIATF-----FPNSAPIHYQGSNPEKEDAQYARMAYKLANWDMQRPPTQE 277

hypothetical protein 032R - Chilo iridescent virus
C;Species: Chilo iridescent virus
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03057
R;Bahr, U.; Tidona, C.A.; Darai, G.
Virus Genes 15, 235-245, 1997
A;Title: The DNA sequence of Chilo iridescent virus between the genome coordinates 0.101
A;Reference number: Z14834; MUID:98141693; PMID:9482589
A;Accession: T03057
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-602 <BAH>
A;Cross-references: UNIPROT:O55720; EMBL:AF003534; NID:G2738385; PIDN:AAB94431.1; PID:G27
C;Superfamily: Chilo iridescent virus hypothetical protein 032R

Query Match 6.9%; Score 122; DB 2; Length 602;
Best Local Similarity 22.1%; Pred. No. 1.5;
Matches 75; Conservative 57; Mismatches 145; Indels 62; Gaps 10;

QY 4 PFGKSHKSPADIVKNLK-----ESMAVLEKQDISDKAEKAT-----REV 43
DB 107 PFLPNFMKPYDLIKNVADPRQKNPFTVLTIEKSGSEKMLTTKSGKPKTFSDMALPEYI 166

QY 44 SKLVAMKEILYGTNEKEPQTEAVAQALAEVLYNSG--LLSTLVADLQIDFEGKKOVAQI 101
DB 167 NSNMTLAEILISINVKKRQTRVTLNEEQIKKASDFKILNSLNLQMLIALIIIAQKLSFI 226

QY 102 -----FNNILRRQIGTRPTVEYICTQONILFMLKGYESPETALNCIMLRCEIRHEPL 156
DB 227 HNDLHFDNVILCKLQRT-FMLVFEVNSICVALLPTYGYPIIIDYGFSEDLIGGPL 285

QY 157 AKIILWSEQYDFRYVEMSTFDIASDAFATKDLLTRHKLKSAFLRHOYDRFSEYK 216
DB 286 LTGIEHNNKGYMNHQY-----DEFTDKTMTLRLSYSGYQGLDKKDAF----- 329

QY 217 LHSENVVTKRQSL---KLGLLELLDRHNTIMTKYISKPENIKLMMNLLRD---KSRNI 270
DB 330 -----OSLIFDKLISKLPIQKQTDGWDKDISVSKQLVRHRIREFVDDYLKSNR 378

QY 271 QFEAFHVKVVFVANPNKTOPILDKLNOAKLIEFLSKF 309
DB 379 E-SFFQKYDYEMVDMTGLSILLPLRKXNTENVETLAIF 416

RESULT 10
D97848
hypothetical protein RC1188 [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: D97848
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: D97848
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-365 <KUR>
A;Cross-references: UNIPROT:Q92GD5; GB:AE006914; PIDN:AAL03726.1; PID:g15620317; GSPDB:
C;Genetics:
A;Gene: RC1188

Query Match 6.6%; Score 117; DB 2; Length 365;
Best Local Similarity 20.1%; Pred. No. 1.7;
Matches 74; Conservative 60; Mismatches 127; Indels 108; Gaps 14;

QY 16 VKNLKESMAVLEKQDISDKAEKATEVSKNIVAMKEILYGTNEKEPQTEAVAQALAEY 75
DB 60 ISLKETAEYTKLHYEVEQSIKARVOEEFLSKQRAIKLAEBEKE-----IAQRIALN 114

QY 76 N--SGLISTLVADLQIDFEGKKDVQAQIFNNILRRQIGTRPTVEYICTQONILFMLKG 133
DB 115 NLHNEFIKNITKOTKRIE-ESNKLDKLINKLENI----- 150

QY 134 YESPEALNCGIMLRECIHHEPIAKIILWSOEYDFFRVEMSTFDIASDAFATPKDLIT 193
DB 151 -----IDHEELNREILTHEIFKLNQ-----AYKVLHK 178
QY 194 RHKLISAEFLEOHYDRFFSEYEKLIHS-----ENVVTKRQSLKLLGELLDRHNFITMT 247
DB 179 NHKKLITEHKQHTB--LNELNKTITNISQLOKGLTPKKVKELGELFEHQEMLKIHK 236
QY 248 KYISKPENLKLMMN---LIRDKSRNIQFEAFHVFKVFVANPNKTOP--ILD-----ILLK 297
DB 237 AHVEKIENSKIVDOEIIKYKEHNINXDKIKKLGSDIKERYKKEPEKYLDAEYKVLK 296
QY 298 NOAKLI-----EFLSKFONDRDTCMS-----SSVPTNSR-----VDLR 331
DB 297 HQHKAIKTGVVKQIEHKNKVLINVDKTKSLANKIREQTKVDKNTSRSLSPARTPNIK 356
QY 332 V-KPRTRGI 339
DB 357 VTNNKTRGI 365
RESULT 11
A:Accession: A90551
C: conserved hypothetical protein MYPU_3130 [imported] - Mycoplasma pulmonis (strain UAB CT)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: A90551
R:Chambaud, I.; Hellig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A905512; MUID:21267165; PMID:11353084
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2819 <KUR>
A:Cross-references: UNIPROT:Q980P8; GB:AL445566; PID:g14089727; PIDN:CAC13486.1; GSPDB:G
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPU_3130
A:Genetic code: SGC3
Query Match 6.5%; Score 115; DB 2; Length 2819;
Best Local Similarity 19.0%; Pred. No. 31;
Matches 83; Conservative 87; Mismatches 127; Indels 140; Gaps 24;
QY 5 FGKSHKSPADIVKXN-----KESMAVLEK-QDISDKAEKATE--EVSKNLVAM 50
DB 1404 FGKT-----NIISNLYDLVLSALINQESSVLTKVNTIHKALDQQNSYFAVITKLLSV 1458
QY 51 KEILYGTNEKEPQTEAVAQAQELY--NSGLLSTLVADL--QLIDFEG--KKDVAQIFNN 104
DB 1459 KE---GDVDSDEFKAIKKFLNFAKNDYLLTKFISPIVDKTIQFDSATKSSIVKLIKN 1515
QY 105 ILRRQ-----IGRTTFVEY-----ICTQONILFM----- 129
DB 1516 VLSDEQNLKPSAQIIDTIFANKNKYLAYENLNDFLDALVDNKENIINSSNTLGKIKTD 1575
QY 130 -----LLKGYESPALNCGIMLRECIHHEPIAKIILWSOEYDFFRVEMSTFDIASDAFATPKDLIT 193
DB 1576 DEFKNIIKSTIKMLMTSDSISITTEINTIVKSAHELIS--IVDSNPLNQLLKDFALL 1633
QY 174 EMSTFDIASDA--FATPKDLITRHKLLSAEFLQHYDRFFSEYEKLIHSENVVTKRQSLK 231
DB 1634 EKNKDKSKTSLEFSTFM-----KXISIEFLDAE--NNYALFKKVLSSNLLSTQS----- 1682
QY 232 LIGELLDRHNF-----IMTKYISKENLKLMMNLLRDKSRNIQ-----EAFHVFKV 280
DB 1683 -----DQGSFTKSISSILRFLRKEKALELM---LKKQPRNFELHDIKVDDAINLIK- 1731
QY 281 FVANPNKTOPITLIDILKKN-----OAKLIEFLSKFONDRDTCMSSSV----- 321
DB 1732 FIFRDNQVIDFETIINRVVSQKEYSKLSYQBIILYKFLSSNKNETINFPKKQIGBIKN 1791

QY 322 -----PTTNSRVDLRVK 333
DB 1792 SKLINPLAKSPIESSELK 1808
RESULT 12
TI8372
repeat organellar protein - Plasmodium chabaudi
C:Species: Plasmodium chabaudi
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: TI8372
R:Werner, E.B.; Taylor, W.R.; Holder, A.A.
Mol. Biochem. Parasitol. 94, 185-196, 1998
A:Title: A Plasmodium chabaudi protein contains a repetitive region with a predicted spe
A:Reference number: Z18922; MUID:98418765; PMID:9747969
A:Accession: TI8372
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1939 <MER>
A:Cross-references: UNIPROT:Q25662; EMBL:U43145; NID:g1151157; PID:g1151158; PIDN:AA0634
Query Match 6.4%; Score 113.5; DB 2; Length 1939;
Best Local Similarity 19.4%; Pred. No. 24;
Matches 57; Conservative 57; Mismatches 113; Indels 67; Gaps 11;
QY 18 NLKESMAVLEKQDISDKAEKATEEVS--KNLVAMKELLYG--TNEKEPQTEAVAQAQAE 73
DB 1433 NLKNVEDKTGDDICEKNNDQAEISYLKDEIKKISMLYGEELNRKNSYDEKVKNLUNE 1492
QY 74 LYNGLLSTLVADLQIDFEGKKDVAQI-----FNNILRRQIGTR 113
DB 1493 L-----KELKIRNKGEEAIAELNKLKNKKNKSVKQNDSSNNIITKD-GDK 1541
QY 114 TPTVEYICTQONI-----LFMLKGYESPALNCGIMLRECIHHEPLAK--IILWSQF 166
DB 1542 TP--EYVSNDDKIQKQKAMLVKLEKPELMDWINSLEKENFRVMSVTKENKQVNDKI 1599
QY 167 YDFRYVEMSTFDIASDAFATPKDLITRHKLLSAEFLQHYDRFFSEYEKLIHSENVYTK 226
DB 1600 VGIYSYFKCEKELKNDMLVI---CLVLKDTLSILFNDNFNFEKIDKILNKQMYIPT 1656
QY 227 -----RQSLKLLGELLDD-----RHNFTIMTKYISKENLKLMM 260
DB 1657 EIRILFLRYSFLDKLRNVYKCVNEVYNNERYEYSWALFQTYLETASNLKEM 1710
RESULT 13
H64574
DNA topoisomerase I - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C:Accession: H64574
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Frazer, C.N
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: H64574
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-677 <TOM>
A:Cross-references: UNIPROT:O25188; GB:AE000559; GB:AE000511; NID:g2133536; PIDN:AA07502
C:Superfamily: DNA topoisomerase I
Query Match 6.3%; Score 112.5; DB 2; Length 677;
Best Local Similarity 19.4%; Pred. No. 7.4;
Matches 84; Conservative 80; Mismatches 141; Indels 129; Gaps 20;
QY 3 FPGKSHKSPADIVKXNKE---SMNVLE--KODISDKK-----AEKATEEVSKNLVAMK 51

Db 222 PKF-KDKNEASQFLKDLKDGAGMSVLSKESLNKEPKPKPTTSLKLSQASKSL----- 276

Qy 52 EILYGNNEKEPOTEAQAQAEALYNSGLLSTLVADLIQDIDFEKGDVAQIFNNILRRQIG 111

Db 277 -----KIP-TKEIAQAQKLUFEAGLIYHRTDSEFLSPYLVKHEVFFPIY----- 322

Qy 112 TRTPTV-----EXIC-----TQONILFMLLKGYESPEIALNCGIMLRECIRH 153

Db 323 ---PSVQYREYKAGKNSQAEEAIRITHPHALKDLEKVCSDAKISEELALKLYLIYT 379

Qy 154 EPI---AKIILWSEQYDFRFRVEMSTFDIA-----SDAFATPKDLITRHKLLSAEFLQ 205

Db 380 NTTCQSQRNALYQ--YDCIFKIKSESFKLSFKLKEKGFLEIBELIQGKEEINREQBS 437

Qy 206 HYDRFF-----SEYEKLIHSENYVTQRQSLKLLGELLL 238

Db 438 EIEFSLKENDSVPLKEVFKKIEKSPKPYKESAFIPLLESBG-IGRSTYASFLLDLL 496

Qy 239 DRHNFITMTK-----YISKPENLKLMMNLLRDKSRNIQFEAFHVFKVFA 283

Db 497 KRKYISIDTKNAITPTSQGLEVISFPKKDKEVDFIALTSDKSK-----LG 543

Qy 284 NPNT-OPILDILLKQAK-----LIEPLSKFOND---RTDCMSSSVPTTNSRVDLRVKP 334

Db 544 NTTQFEEDLLIMRGEASYEKFMLEVISKLSTAKFYQYQSTDNMNPDK---QLELID 600

Qy 335 RTRGIRDLPKPAQ 348

Db 601 KICKDKKLQPSOE 614

RESULT 14

T18489

hypothetical protein C0820w - malaria parasite (Plasmodium falciparum)

C;Species: Plasmodium falciparum

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T18489

R;Lawson, D.; Bowman, S.; Barrell, B.

submitted to the EMBL Data Library, August 1997

A;Reference number: Z18935

A;Accession: T18489

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-4981 <LAW>

A;Cross-references: UNIPROT:O77372; EMBL:Z98551; NID:e1331910; PIDN:CAB111

C;Genetics:

A;Map position: 3

A;Note: C0820w

Query Match 6.2%; Score 111; DB 2; Length 4981;

Best Local Similarity 20.1%; Pred. No. 1.2e+02;

Matches 59; Conservative 49; Mismatches 104; Indels 82; Gaps 10;

Qy 100 QIFNNILRRQIGTRTPVYICTQONILFMLLKYES-----PEIALNCGIMLRECIRHP 155

Db 1521 KIFLNYIRKEKEHIYNFQYITTFPTFLEMLKNIQSDKDYKFTLYNIGIINMMCVAVKF 1580

Qy 156 LA-----KII-----LWSEQ-----FYDFRFRVEMSTFDIASDAFATFKOLL 192

Db 1581 LKEKCPNKIISIFFKLLAEVPIGEEEDKVVTLYKKNYSYIYRTHF-----FNLYKDKC 1634

Qy 193 TRHKLLSAEFLQHYDRFFSEYEKLIHSENYVTQRQSLKLLGELLLDRHNFITMTKYIS- 251

Db 1635 T-----TLYTKNCMEYIKWMRNINWLKKKN-----DSQVDFNFIILSNMYN 1678

Qy 252 -----KPENLKLMMNLLRDKSRNIQFEAFHVFKVFN-----PNKT 288

Db 1679 SDTFKKXNIQVLYNDYQQYVKKNVVHIENMYSVVVTHKEMENGMYRKEGKRENT 1738

Qy 289 QPILDILLKNO-----AKLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVK 333

Db 1739 YDKINIDDKNFGQPIYKYVVDKLTSKRNTLNDSSINCLLESSYVTKGNTKIRIK 1792

RESULT 15

G70163

hypothetical protein BB0512 - Lyme disease spirochete

C;Species: Borrelia burgdorferi (Lyme disease spirochete)

C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

C;Accession: G70163

R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whittson, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A;Authors: Smith, H.O.; Venter, J.C.

A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A;Reference number: A70100; MUID:98065943; PMID:9403685

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-2166 <KL8>

A;Cross-references: UNIPROT:O51465; GB:AE001153; GB:AE000783; NID:g2688419; PIDN:AAC6687

A;Experimental source: strain B31

Query Match 6.1%; Score 109; DB 2; Length 2166;

Best Local Similarity 22.5%; Pred. No. 55;

Matches 80; Conservative 68; Mismatches 114; Indels 94; Gaps 23;

Qy 17 KNLKES--MAVLEKQDISDKAEKATEEYSK---NLVAMKEILYGTNEKEPQTEAVAQLA 71

Db 1623 KNIGESLNKLVLDLEKXVDFKLEKIDKVKNKTKTDILIOAEVFLTQCKDLE-DKIFELN 1681

Qy 72 QELYNSGLLSTLVADL-----QLID-----FEKG-----KDVQAQIFNNI--LRRQIG 111

Db 1682 QKLEHE--FTYLSNLDKVRREWVDVISSDKSEFQEGIELINKNISFSEKISLYRNIE 1739

Qy 112 TRTPTVYICTQONILFMLLKGYESPEIALNCGIM---LRECIRH-----EP 155

Db 1740 TSIEH-EY-----NSFSKSIKSDGLLEDELKSLKSHSTSEITIKSGHQEQ 1785

Qy 156 LAKI-ILWSEQYDFRFRVEMSTFDIASDA-----FATFKDLITRHKL-LSAEFL 203

Db 1786 IDRFEVEFKNKHLEKVDNILESKILNCVQPNKFISEIKDNLVYKSDLRAEF- 1844

Qy 204 EQHYDRFFSEYEKLIHSENYVTQRQSLKLLGELLLDRHNFITMTKYISKP-----ENLK 257

Db 1845 EDSYDKINFOIENQI--ENF-----KCLDSE--LEKNNSIFLEAYSCLKLEKLEWETLK 1894

Qy 258 LMNMLLRDKSRNIQFEAFHVFKVFNPNKTPQIPILDILLKNOAKL-IEFLSKFOND 312

Db 1895 NEIGLAQEQYKNFE---NVNKEFYNTQKETLGIIEIF--NELKLEQESIKSIKND 1944

Search completed: October 22, 2004, 09:13:07

Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2004, 08:56:03 ; Search time 191 Seconds
(without alignments)
1054.351 Million cell updates/sec

Title: US-10-089-688-2

Perfect score: 1781

Sequence: 1 MPFFPGKHKSPADIVKNLK.....RVKPTRGIRDLKRPAAQQA 350

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1635.5	91.8	341	1	CB39_HUMAN	Q9Y376 homo sapien
2	1619.5	90.9	341	2	Q8VDZ8	Q8VDZ8 mus musculus
3	1614.5	90.7	341	1	CB39_MOUSE	Q61P72 xenopus lae
4	1614.5	90.7	341	2	Q61P72	Aah72045 xenopus l
5	1614.5	90.7	341	2	AAH72045	Q74X00 homo sapien
6	1590.5	89.3	341	2	Q724X0	Q61Q12 brachydanio
7	1537.5	86.3	341	2	Q61Q12	Aah71393 brachydanio
8	1537.5	86.3	341	2	AAH71393	Q6PAB4 xenopus lae
9	1339.5	75.2	337	2	Q6PAB4	Q9db16 mus musculus
10	1339.5	75.2	337	2	AAH60384	Q9H9S4 homo sapien
11	1298.5	72.9	337	1	C39L_MOUSE	Aaq3064 homo sapi
12	1297.5	72.9	337	1	C39L_HUMAN	Q803V8 brachydanio
13	1297.5	72.3	337	2	AAQ3064	Q7P07 anopheles g
14	1287	72.3	343	2	Q803V8	Q7P05 anopheles g
15	1163.5	65.3	332	2	Q7P07	P91891 drosophila
16	1163.5	65.3	335	2	Q7P8N5	Aam75031 drosophil
17	1123	63.1	339	1	MO25_DROME	Q722A5 caenorhabdi
18	1123	63.1	339	2	AAH75031	Q12643 caenorhabdi
19	1054.5	59.2	377	2	Q722A5	O18211 caenorhabdi
20	1054.5	59.2	336	2	Q21643	Q8K312 mus musculus
21	1006	56.5	338	1	MO2M_CAEEL	Q9Yf6 chlorella p
22	944.5	53.0	205	2	Q8K312	Q9f6g3 arabidopsis
23	766	43.0	329	1	YFV6_SCHPO	Aal16128 arabidops
24	764.5	42.9	321	1	DE76_CHLPR	Q9m0m4 arabidops
25	726	40.8	390	2	Q6CAN7	Aal06959 arabidops
26	711.5	39.9	343	1	MO2N_ARATH	Q6BSQ5 debaryomyce
27	711.5	39.9	343	1	AAL16128	Q81919 arabidopsis
28	705.5	39.6	343	1	MO2M_ARATH	
29	705.5	39.6	343	2	QAL05959	
30	686	38.5	338	2	Q6BSQ5	
31	669.5	37.6	345	2	Q81919	

32 651.5 36.6 337 2 Q7XIQ4
33 623 35.0 384 1 HYMA_EMENI
34 614 34.5 348 1 MO2L_ARATH
35 614 34.5 348 2 AAR24657
36 584 32.8 370 2 Q873K5
37 571 32.1 364 2 Q6CWP5
38 560 31.4 362 2 Q757C0
39 560 31.4 362 2 AAS52777
40 511 28.7 422 2 Q6FMG7
41 501 28.1 399 1 HYN1_YEAST
42 455.5 25.6 332 2 Q7OTG6
43 327 18.4 103 2 Q8K038
44 283.5 15.9 509 2 Q7YVL6
45 137.5 7.7 339 1 MO2L_CAEEL

Q7xiq4 oryza sativ
O60032 emericella
Q9zq77 arabidopsis
Aar24657 arabidops
Q873k5 neurospora
Q6cwp5 kluyveromyc
Q757c0 ashbya goss
Aas52777 ashbya go
Q6fmw7 candida gla
F32464 saccharomyc
Q7otg6 giardia lam
Q8k038 mus musculu
Q7yvl6 cryptospori
Q9tzm2 caenorhabdi

ALIGNMENTS

RESULT 1
CB39_HUMAN
ID CB39_HUMAN STANDARD; PRT; 341 AA.
AC Q9Y376;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Calcium binding protein 39 (Mo25 protein) (CGI-66).
GN Name=CB39; Synonyms=MO25;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20272150; PubMed=10810093;
RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
RT "Identification of novel human genes evolutionarily conserved in
RT Caenorhabditis elegans by comparative proteomics.";
RL Genome Res. 10:703-713(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RA Jin W., Shi J., Ren S., Gu J., Fu S., Huang Q., Dong H., Yu Y., Fu G.,
RA Wang Y., Chen Z., Han Z.;
RT "A novel gene expressed in the human hypothalamus";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Duoenum;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci E., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Maman A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
Rach G., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.85 ÅNGSTROMS) OF 10-339.
RX PubMed=14730349; DOI=10.1038/nsmb716;

RA Milburn C.C., Boudeau J., Deak M., Alessi D.R., van Aalten D.M.;
RT "Crystal structure of Mo25 alpha in complex with the C terminus of the
RL pseudo kinase STE20-related adaptor.";
RL Nat. Struct. Mol. Biol. 11:193-200(2004).
CC -I- FUNCTION: Together with the STE20-related adaptor-alpha (STRAD
CC alpha) pseudo kinase, forms a regulatory complex capable of
CC stimulating the activity of STR11.
CC
CC -I- SIMILARITY: Belongs to the Mo25 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF151824; AAD34061.1; -
CC EMBL; AF113536; AAF14873.1; -
CC EMBL; BC020570; AAH20570.1; -
CC DR PDB; 1UPK; X-ray; A=1-341.
CC DR IntAct; Q9Y376; -
CC DR Genew; HGNC:20292; CAB39.
CC DR InterPro; IPR008938; ARM.
CC DR InterPro; IPR004892; Mo25.
CC DR Pfam; PF03204; Mo25; 1.
CC KW 3D-structure.
CC SQ SEQUENCE 341 AA; 39869 MW; EC710A528B6F9811 CRC64;

Query Match 91.8%; Score 1635.5; DB 1; Length 341;
Best Local Similarity 94.0%; Pred. No. 1e-96;
Matches 329; Conservative 1; Mismatches 11; Indels 9; Gaps 2;

QY 1 MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
DB 1 MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
QY 61 EPQTEAVAQAELYNGLSTLVADLQIDFEGKDVQAFNNILRRQIGTRPTVEYI 120
DB 61 EPQTEAVAQAELYNGLSTLVADLQIDFEGKDVQAFNNILRRQIGTRPTVEYI 120
QY 121 CTQONILFMLLKGVESPEIALNCGIMLRCEIRHEPLAKIILWSQYDFFRYVEMSTFDI 180
DB 121 CTQONILFMLLKGVESPEIALNCGIMLRCEIRHEPLAKIILWSQYDFFRYVEMSTFDI 180
QY 181 ASDAFATFKDLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKQSLKLGELLDR 240
DB 181 ASDAFATFKDLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKQSLKLGELLDR 240
QY 241 HNFITMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQA 300
DB 241 HNFITMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQA 300
QY 301 KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVKPRTRGIRDLKPAQOE 350
DB 301 KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVKPRTRGIRDLKPAQOE 341

RESULT 2
Q8VDZ8 PRELIMINARY; PRT; 341 AA.
AC Q8VDZ8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cab39 protein.
OS Mus musculus (Mouse).
GN Name=Cab39;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=CZECH II;
RX TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Shevchenko Y., Bouffard G.G.,
RA Whiting M., Madan A., Young A.C., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Schmutz J., Schein J.E.,
RA Krzywinski M.I., Skalska U., Smalusz D.B., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL [2].
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II;
RC TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
RA Strausberg R.;
RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC020041; AAH20041.1; -
DR MGD; MGI:107438; Cab39.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
DR SQ SEQUENCE 341 AA; 39843 MW; E7FECAS29D6FE811 CRC64;

Query Match 90.9%; Score 1619.5; DB 2; Length 341;
Best Local Similarity 93.1%; Pred. No. 1.1e-95;
Matches 326; Conservative 1; Mismatches 14; Indels 9; Gaps 2;

QY 1 MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
DB 1 MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
QY 61 EPQTEAVAQAELYNGLSTLVADLQIDFEGKDVQAFNNILRRQIGTRPTVEYI 120
DB 61 EPQTEAVAQAELYNGLSTLVADLQIDFEGKDVQAFNNILRRQIGTRPTVEYI 120
QY 121 CTQONILFMLLKGVESPEIALNCGIMLRCEIRHEPLAKIILWSQYDFFRYVEMSTFDI 180
DB 121 CTQONILFMLLKGVESPEIALNCGIMLRCEIRHEPLAKIILWSQYDFFRYVEMSTFDI 180
QY 181 ASDAFATFKDLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKQSLKLGELLDR 240
DB 181 ASDAFATFKDLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKQSLKLGELLDR 240
QY 241 HNFITMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQA 300
DB 241 HNFITMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQA 300
QY 301 KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVKPRTRGIRDLKPAQOE 350
DB 301 KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVKPRTRGIRDLKPAQOE 341

RESULT 3
CB39_MOUSE
ID CB39_MOUSE STANDARD; PRT; 341 AA.
AC Q06138;
DT 01-FEB-1994 (Rel. 26, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Calcium binding protein 39 (Mo25 protein).

GN Name=Cab39; Synonyms=Mo25;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93119656; PubMed=8418809;
 RA Miyamoto H., Matsushiro A., Nozaki M.;
 RT "Molecular cloning of a novel mRNA sequence expressed in cleavage
 RL stage mouse embryos";
 CC Mol. Reprod. Dev. 34:1-7(1993).
 CC -!- FUNCTION: Together with the STE20-related adaptor-alpha (STRAD
 CC alpha) pseudo kinase, forms a regulatory complex capable of
 CC stimulating the activity of SRK1.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -!- DEVELOPMENTAL STAGE: Transcribed during early mouse development.
 CC Detected at all developmental stages from the egg through the
 CC blastocyst, most abundant at the 2-cell stage.
 CC -!- SIMILARITY: Belongs to the Mo25 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; S51858; AAB24801.1; -;
 DR MGD; MGI:107438; Cab39.
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR004892; Mo25.
 DR Pfam; PF03204; Mo25; 1.
 SQ SEQUENCE 341 AA; 39842 MW; E7F668529D6FE811 CRC64;

Query Match 90.7%; Score 1614.5; DB 1; Length 341;
 Best Local Similarity 92.9%; Pred. No. 2.2e-95;
 Matches 325; Conservative 2; Mismatches 14; Indels 9; Gaps 2;

QY 1 MPFFPGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
 DB 1 MPFFPGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
 QY 61 EPQTEAVAQLAQELYNGLSTLVADLQIDFEGKDVQAFNNILRRQIGTRPTVEYI 120
 DB 61 EPQTEAVAQLAQELYNGLSTLVADLQIDFEGKDVQAFNNILRRQIGTRPTVEYI 120
 QY 121 CTQQNILFMLLKGYSPFIALNCGIMLRECIHREPLAKIILWSEQFYDFRYVEMSTFDI 180
 DB 121 CTQQNILFMLLKGYSPFIALNCGIMLRECIHREPLAKIILWSEQFYDFRYVEMSTFDI 180
 QY 181 ASDAFATFKDILLTRHKLSSAEFLQHYDRFFSEYEKLLHSENYVTKQSLKLGELLDR 240
 DB 181 ASDAFATFKDILLTRHKLSSAEFLQHYDRFFSEYEKLLHSENYVTKQSLKLGELLDR 240
 QY 241 HNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNOA 300
 DB 241 HNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNOA 300
 QY 301 KLIEFLSKFQNDRTDCSSVPTTNSRVLDLRKVRTRGIRDLKRAQOE 350
 DB 301 KLIEFLSKFQNDRTDCSSVPTTNSRVLDLRKVRTRGIRDLKRAQOE 350
 DB 301 KLIEFLSKFQNDRTDCSSVPTTNSRVLDLRKVRTRGIRDLKRAQOE 350

RESULT 4
 Q6IP72 PRELIMINARY; PRT; 341 AA.
 AC Q6IP72;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE MGC78903 protein.

GN Name=MGC78903;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=89355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kertman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative";
 RT Dev. Dyn. 225:384-391 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC072045; AAH72045.1; -;
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR004892; Mo25.
 DR Pfam; PF03204; Mo25; 1.
 SQ SEQUENCE 341 AA; 39777 MW; 5AFAF9325BB2F16 CRC64;

Query Match 90.7%; Score 1614.5; DB 2; Length 341;
 Best Local Similarity 92.3%; Pred. No. 2.2e-95;
 Matches 323; Conservative 5; Mismatches 13; Indels 9; Gaps 2;

QY 1 MPFFPGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
 DB 1 MPFFPGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGTNEK 60
 QY 61 EPQTEAVAQLAQELYNGLSTLVADLQIDFEGKDVQAFNNILRRQIGTRPTVEYI 120
 DB 61 EPQTEAVAQLAQELYNGLSTLVADLQIDFEGKDVQAFNNILRRQIGTRPTVEYI 120
 QY 121 CTQQNILFMLLKGYSPFIALNCGIMLRECIHREPLAKIILWSEQFYDFRYVEMSTFDI 180
 DB 121 CTQQNILFMLLKGYSPFIALNCGIMLRECIHREPLAKIILWSEQFYDFRYVEMSTFDI 180
 QY 181 ASDAFATFKDILLTRHKLSSAEFLQHYDRFFSEYEKLLHSENYVTKQSLKLGELLDR 240
 DB 181 ASDAFATFKDILLTRHKLSSAEFLQHYDRFFSEYEKLLHSENYVTKQSLKLGELLDR 240
 QY 241 HNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNOA 300
 DB 241 HNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNOA 300

QY	301	KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVKPRTGIRDLKRPAAQOEA	350			
Db	301	KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVKPRTGIRDLKRPAAQOEA	341			
RESULT 5						
AAH72045						
ID	AAH72045	PRELIMINARY;	PRT; 341 AA.			
AC	AAH72045;					
DT	01-JUN-2004	(TEMBLrel. 27, Created)				
DT	01-JUN-2004	(TEMBLrel. 27, Last sequence update)				
DT	01-JUN-2004	(TEMBLrel. 27, Last annotation update)				
DE	Hypothetical protein.					
OS	Xenopus laevis (African clawed frog).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;					
OC	Xenopodinae; Xenopus; Xenopus.					
OX	NCBI_TaxID=8355;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Spleen;					
RX	MEDLINE=22341132; PubMed=12454917;					
RA	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,					
RA	Richardson P.;					
RT	"Genetic and genomic tools for Xenopus research: The NIH Xenopus					
RT	initiative.";					
RL	Dev. Dyn. 225:384-391 (2002).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Spleen;					
RX	MEDLINE=22388257; PubMed=12477932;					
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,					
RA	Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,					
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,					
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,					
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,					
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,					
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,					
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,					
RA	Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,					
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,					
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,					
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,					
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,					
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,					
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,					
RA	Krzywinski M.I., Skalska U., Smallus D.E., Schmerch A., Schein J.E.,					
RA	Jones S.J., Marra M.A.;					
RT	"Generation and initial analysis of more than 15,000 full-length human					
RT	and mouse cDNA sequences.";					
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).					
RN	[3]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Spleen;					
RA	Klein S., Strausberg R.;					
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.					
DR	EMBL; BC072045; AAH72045.1; -.					
KW	Hypothetical protein.					
SQ	SEQUENCE 341 AA; 39777 MW; 5AFAF9325B2F16 CRC64;					
Query Match 90.7%; Score 1614.5; DB 2; Length 341;						
Best Local Similarity 92.3%; Pred. No. 2.2e-95;						
Matches 323; Conservative 5; Mismatches 13; Indels 9; Gaps 2;						
QY	1	MPFPFGSHKSPADIVKNLKESMAVLEKQDISDKAKATEEVSKNLVAMKEILYGTNEK	60			
Db	1	MPFPFGSHKSPADIVKNLKESIAVLEKQDISDKAKATEEVSKNLVAMKEILYGTNEK	60			
QY	61	EPQTEAQAQLAQLYNSGLSTLVADLQIDFEGKKDVAQIFNNILRRQIGTRTPTVEYI	120			
Db	61	EPQTEAQAQLAQLYNSGLTGLVADLQIDFEGKKDVAQIFNNILRRQIGTRTPTVEYI	120			

05-JUL-2004 (TReMBLrel. 27, Last annotation update)
 Zgc:86716.
 Name=zgc:86716;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_TaxID=7955;
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Embryo;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC071393; AAH71393.1; -.
 DR InterPro; IPR008938; ARM.
 DR Pfam; PF03204; Mo25; 1.
 SQ SEQUENCE 341 AA; 39792 MW; 9CEE4A665F3C6614 CRC64;

 Query Match 86.3%; Score 1537.5; DB 2; Length 341;
 Best Local Similarity 87.4%; Pred. No. 1.9e-90;
 Matches 306; Conservative 14; Mismatches 21; Indels 9; Gaps 1;

 Qy 1 MPFPFGKSHKSPADIVKNLKESMAVLEKODISDKAEKATEEVSKNLVAMKEILYGTNEK 60
 Db 1 MPFPFVKSHKCPADIVKNLKNMTILEKODISDKAEKASEVSKLSMKELLYGTNEK 60

 Qy 61 EPQTEAVAQAQLYNSGLLSTLVADLQIDFEGKQVQAFNNILRRQIGTRPTVEYI 120
 Db 61 EPQTEAVAQAQLYNSGLLSTLVADLQIDFEGKQVQAFNNILRRQIGTRPTVEYL 120

 Qy 121 CTQONILFMLKGVESPIALNCGIMLRECIHREPLAKIILWSQYDFFRVEMSTFDI 180
 Db 121 CTQONILFMLKGVESPIALNCGIMLRECIHREPLAKITLCSQYDFFRVEMSTFDI 180

 Qy 181 ASDAFATFKOLLTRHKLLSAEFLQHYDRFPFSEYKLLHSENYVTKQSLKLGELLDR 240
 Db 181 ASDAFATFKOLLTRHKLLSAEFLQHYDRFPFSEYKLLHSENYVTKQSLKLGELLDR 240

 Qy 241 HNFITMTKYISKPNLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQIPDILLKNQA 300
 Db 241 HNFITMTKYISKPNLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPPILDILLKNQT 300

 Qy 301 KLIFLSKFQNDRTDCMSSSVPTTNSRDLRVKPRTRGIRDLKRPQAOEA 350
 Db 301 KLIFLSKFQNDRAEDEQFSDEKTY-----LIKQIRDLKRPQDA 341

120 SQSHQILFFILKGVSPQVALHCGIMLRQCVRHPLAKVILYSQFGDPFKYVEMSTEDI 177

181 ASDAFATFKDLTRHKLKLSAEFLFQHYDRFPSEYKELLHSENVTYKQSLKLGELLDR 240

180 ASDAFATFKDLTRHKLKLSAEFLFQHYDRFPSEYKELLHSENVTYKQSLKLGELLDR 239

241 HNTFTMTKYISKPENIKLMMNLKDKSRNTOPEAFHVFVKVFNPNKTPQILDILKNOA 300

240 HNFSTMTKYISKPENIKLMMNLKDKSPNIOQFAHVFVKVFNPNKTPQILDILKNOA 299

301 KLIEFLSKFQNDRTDCMSSSVPTTNSRDLRVKPRTRGRDILKRP 345

300 KLIDFLSKFQNDRTDCMSSSVPTTNSRDLRVKPRTRGRDILKRP 335

RESULT 10

AAH60384

ID AAH60384 PRELIMINARY; PRT; 337 AA.

AC AAH60384;

DT 02-MAR-2004 (TReMBLrel. 27, Created)

DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)

DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)

GN MGC68674 protein.

DE MGC68674.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

OC Xenopodidae; Xenopus; Xenopus.

NCBI_TaxID=8355;

[1]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RX MEDLINE=22341122; PubMed=12454917;

RA Klein S.B., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,

RA Richardson P.;

RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus

RT initiative.";

RL Dev. Dyn. 225:384-391(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Sherman C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Datschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RA Klein S., Strausberg R.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.

RA EMBL; BC060384; AAH60384.1; --

SQ SEQUENCE 337 AA; 39183 MW; CCA5608F5D856957 CRC64;

Qy	1	MPFPFGKSHKSPADIVKNIKESMAVLEKQDISDKKAEKATEEVSKNILVAMKEILYGTNEK	60
Db	4	MPL-FSKSHKPAEIVKILKUNMALLERQ---DKTERASEEVSLSQATKEILCTGDK	59
Qy	61	BPQTEAVAQLAQELYSGLLSTLVADQLIDPEGKKDVAQIENNILRRROIQTRTPTVEY	120
Db	60	EPQETVAQLAQELYSGLLVLIANLHLIDPEGKKDVSQIENNILRRQIGTRSPITVEY	119
Qy	121	CTQQNILFWLKGYSPEIALNCGIMLRCTIRHEPLAKIILNBSQYDFFRYVEMSTFDI	180
Db	120	SSHQILFIILKGYESPQVHALHCGIMLRCEVHEPLAKVILYSEQDGFVKYVEMSTFDI	179
Qy	181	ASDAFATKDLLTRKKLSAEPLQEHYORFPSEYEKLLHSENVYTKQSLKILGELLDR	240
Db	180	ASDAFATKDLLTRHKLMVAEPLQYDRIFNDYEKLLHSENVYTKQSLKILGELLDR	239
Qy	241	HNFTIMTKYIISKPENLKLMMNILLRDKSRNIQPEAFHFVKVFVANPNKTOPIIDLILKNQA	300
Db	240	HNFSIMTKYIISKPENLKLMMNILLRDKSPNIQPEAFHFVKVFVANPNKTOPIVDILLKNQT	299
Qy	301	KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLVRVKPRTGRIGRIDUKRP	345
Db	300	KLIDFLSSFKQDRDT---DEQFTDEKNYL-----IKOIRDUKKP	335

RESULT 17

Absolu
RN _ II
C39L MOUSE
ID -C39L MOUSE STANDARD; PRT; 337 AA.
AC Q9DBI6; O8BG52; Q91WB8 ; Q91YL0;
AD DT 16-OCT-2001 (Rel. 40, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Calcium binding protein 39-like (Mo25-like protein).
GN Cab39l; GN
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NBIB TaxID=10090;
RN [1]

RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC	STRAIN=C57BL/6J; TISSUE=Cerebellum, Eye, Pituitary, and Testis;
RX	MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA	Nikaido I., Osato K., Saito R., Suzuki H., Yamanaka I., Kiyoosawa H.,
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,
RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA	Schriml L.M., Kanapin A., Matsuda H., Batalov S., Belsel K.W.,
RA	Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA	Dalla E., Draganí T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA	Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA	Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA	Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,
RA	Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA	Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA	Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA	Petrovsky N., Pallat R., Pontius J.U., Qi D., Ramachandran S.,
RA	Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA	Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA	Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA	Verardo R., Wagner L., Wahlstedt C., Wang Y., Matanabe Y., Wells C.,
RA	Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA	Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA	Hiroxane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA	Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA	Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA	Miyazaki A., Sakai K., Saeki D., Shibata K., Shinagawa A.,
RA	Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA	Birney E., Hayashizaki Y.
RT	"Analysis of the mouse transcriptome based on functional annotation of
RT	60,770 full-length cDNAs."
RL	Nature 420:563-573(2002).
RN	[2]

121 CTTQOILFLLKGYSPETALNCGIMLRCEIRHEPLAKIILMSQYDFRFRVYVMTFPI 180
 120 SSSPHILFLLKGYEAPQIALRCGIMLRCEIRHEPLAKIILMSQYDFRFRVYVMTFPI 179
 181 ASDAFATFOLLTRHKLKLSAEFLQHYDFFSEYKLLHSENVYVTKRSIKLLGELLDR 240
 180 ASDAFATFOLLTRHKLKLSAEFLQHYDFFSEYKLLHSENVYVTKRSIKLLGELLDR 239
 241 HNTTMTKYSKPENIKLMMNLLRDKSRNIQPEAFHVFVFNPNKTPILDILLKNOA 300
 240 HNTTMTKYSKPENIKLMMNLLRDKSRNIQPEAFHVFVFNPNKTPILDILLKNOA 299
 301 KLIEFLSKFQNDRTQWSSSVPTNSRVLDVPRTRGIRDLKRP 346
 300 KLIEFLSSQKERTD-----DEQFADEKNYLIKQIRDLKAA 336
 RESULT 12
 C39L_HUMAN STANDARD; PRT; 337 AA.
 ID C39L_HUMAN PRT; 337 AA.
 AC Q9H954; Q9H954; Q9B233;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 01-OCT-2004 (Rel. 45, Last sequence update)
 DE 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Calcium binding protein 39-like (Mo25-like protein) (Antigen MAA-34).
 GN Name=CAB39L;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chen G., Zhang W.G., Fu J., Cao X.M., Chen Y.X., Zhao W.H., Zhao A.Z.,
 RA Han Y.H., Li P.Y., Liu X.P., Yao L.B.;
 RT "Serological identification of immunogenic antigens in acute monocytic
 RT leukemia";
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [3]
 RP SEQUENCE FROM N.A.
 RA Duhamel A., Matthews L.H., Burton J., Ashurst J.L., Howe K.L.,
 RA Ashcroft K.J., Beare D.M., Burford D.C., Hunt S.E., Scott C.E.,
 RA Griffiths-Jones S., Jones M.C., Keenan S.J., Oliver K., Scott C.E.,
 RA Ainscough R., Almeida J.P., Ambrose K.D., Andrews D.T.,
 RA Aswell R.I.S., Babbage A.K., Baguley C.L., Bailey J., Bannerjee R.,
 RA Barlow K.F., Bates K., Beasley H., Bird C.P., Bray-Allen S.,
 RA Brown A.J., Brown J.Y., Burrell W., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M.E., Clark S.Y., Clarke G., Clee C.M.,
 RA Clegg S.C., Cobley V., Collins J.E., Corby N., Coville G.J.,

Deloukas P., Dhami P., Dunham I., Dunn M., Earthrowl M.E., French L.,
 RA Ellington A.G., Faulkner L., Frankish A.G., Frankland J., French L.,
 RA Garner P., Garnett J., Gilbert J.G.R., Gilson C.J., Ghori J.,
 RA Grafham D.V., Griffiths S.M., Griffiths C., Hall R.E., Hammond S.,
 RA Harley J.L., Hart E.A., Heath P.D., Howden P.J., Huckle E.J.,
 RA Hunt P.J., Hunt A.R., Johnson C., Johnson D., Kay M., Kimberley A.M.,
 RA King A., Laird G.K., Langford C.J., Lawlor S., Leongamornlert D.A.,
 RA Lloyd D.M., Lloyd C., Loveland J.E., Lovell J., Martin S.,
 RA Mashreghi-Mohammadi M., McLaren S.J., McMurray A., Milne S.,
 RA Moore M.J.F., Nickerson T., Palmer S.A., Pearce A.V., Peck A.I.,
 RA Pelan S., Phillimore B., Porter K.M., Rice C.M., Searle S.,
 RA Sehra H.K., Showkeen R., Skuce C.D., Smith M., Steward C.A.,
 RA Sycamore N., Tester J., Thomas D.W., Tracey A., Tromans A., Tubby B.,
 RA Wall M., Wallis J.M., West A.P., Whitehead S.L., Willey D.L., Tubby B.,
 RA Wilming L., Wray P.W., Wright M.W., Young L., Coulson A., Durbin R.,
 RA Hubbard T., Sulston J.E., Beck S., Bentley D.R., Rogers J., Ross M.T.;
 RT "The DNA sequence and analysis of human chromosome 13";
 RL Nature 428:522-528 (2004).
 [4]
 RP SEQUENCE OF 7-337 FROM N.A.
 RX PubMed=14702039; DOI=10.1038/ngl285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori K., Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama K., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Komatsu T.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isoqai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs";
 RL Nat. Genet. 36:40-45 (2004).
 CC -1- SIMILARITY: Belongs to the Mo25 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AY288977; AAQ93064.1; -
 CC EMBL; BC010993; AAH10993.2; -
 CC EMBL; AL138875; CAC28084.1; -
 CC EMBL; AK022639; BAB14147.1; ALT_INIT.
 CC Genew; HGNC:20290; CAB39L.
 CC InterPro; IPR008938; ARM.
 CC InterPro; IPR004892; Mo25.
 CC Pfam; PF03204; Mo25; 1.
 CC CONFLICT 58 58
 CC CONFLICT 93 93
 CC CONFLICT 113 113
 CC E -> D (in Ref. 1).
 CC G -> E (in Ref. 2).
 CC S -> C (in Ref. 1).


```

FT CONFLICT 116 116 V -> A (in Ref. 1).
FT CONFLICT 121 121 A -> S (in Ref. 1).
FT CONFLICT 127 127 F -> V (in Ref. 1).
FT CONFLICT 156 156 A -> V (in Ref. 4).
FT CONFLICT 243 243 A -> T (in Ref. 1).
FT CONFLICT 335 335 T -> A (in Ref. 1).
SQ SEQUENCE 337 AA; 39087 MW; 5C6C7A83B9B4CA2D CRC64;

Query Match
Best Local Similarity 72.9%; Score 1297.5; DB 1; Length 337;
Matches 263; Conservative 32; Mismatches 38; Indels 13; Gaps 3;

QY 1 MPFPFGKSHKSPADIVKNLKSMAVLEKQDISDKKAEKATEEVSQNLVAMKEILYGTNEK 60
DB 4 MPL-FSKSHKSPADIVKNLKSMAVLEKQDISDKKAEKATEEVSQNLVAMKEILYGTNEK 59
QY 61 EPQTEAVAQAQLYNSGLLSTLVADLQIDFEGKDVQVAFNNILRRQIGTRTPTVEYI 120
DB 60 EPPTTEAVAQAQLYSSGLLVTLADLQIDFEGKDVQVAFNNILRRQIGTRTPTVEYI 119
QY 121 CTQONILFMLKGYBSPETALNCGIMLRCEIRHEPLAKIILWSEQYDFFRYVEMSTFDI 180
DB 120 SAHPHILEFLLKGYBAPQIALRCGIMLRCEIRHEPLAKIILWSEQYDFFRYVEMSTFDI 179
QY 181 ASDAFATKOLLTRHKLLSABELEGHYDRFPFSEYEKLLHSENYVTKRQSLKLLGELLDR 240
DB 180 ASDAFATKOLLTRHKLLSABELEGHYDRFPFSEYEKLLHSENYVTKRQSLKLLGELLDR 239
QY 241 HNFITMTKYISKPENLKMMLLRDKSRNIOPEAFHFVKFVFNPNKTOPILDILLKNOA 300
DB 240 HNFALMTKYISKPENLKMMLLRDKSRNIOPEAFHFVKFVFNPNKTOPILDILLKNOA 299
QY 301 KLIEFLSFQNDRTDCMSSVPTTNSRVDLRVKPRTRGIRDILKRP 346
DB 300 KLIEFLSSFOKERTD-----DEQFADEKNYLIKQIRDLKKA 336

RESULT 13
AAQ93064 ID AAQ93064 PRELIMINARY; PRT; 337 AA.
AC AAQ93064;
DT 31-MAR-2004 (TREMELrel. 27, Created)
DT 31-MAR-2004 (TREMELrel. 27, Last sequence update)
DT 31-MAR-2004 (TREMELrel. 27, Last annotation update)
DE Antigen MUA-34.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen G., Zhang W.G., Fu J., Cao X.M., Chen Y.X., Zhao W.H., Zhao A.Z.,
RA Han Y.H., Li F.Y., Liu X.P., Yao L.B.;
RA "Serological identification of immunogenic antigens in acute monocytic
RA leukemia";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY288977; AAQ93064.1; -.
SQ SEQUENCE 337 AA; 39029 MW; B8F1AB5C7905EC93 CRC64;

Query Match
Best Local Similarity 72.3%; Score 1287.5; DB 2; Length 337;
Matches 261; Conservative 33; Mismatches 39; Indels 13; Gaps 3;

QY 1 MPFPFGKSHKSPADIVKNLKSMAVLEKQDISDKKAEKATEEVSQNLVAMKEILYGTNEK 60
DB 4 MPL-FSKSHKSPADIVKNLKSMAVLEKQDISDKKAEKATEEVSQNLVAMKEILYGTNEK 59
QY 61 EPQTEAVAQAQLYNSGLLSTLVADLQIDFEGKDVQVAFNNILRRQIGTRTPTVEYI 120
DB 60 EPPTTEAVAQAQLYSSGLLVTLADLQIDFEGKDVQVAFNNILRRQIGTRTPTVEYI 119
QY 121 CTQONILFMLKGYBSPETALNCGIMLRCEIRHEPLAKIILWSEQYDFFRYVEMSTFDI 180
DB 120 SAHPHILEFLLKGYBAPQIALRCGIMLRCEIRHEPLAKIILWSEQYDFFRYVEMSTFDI 179

```

```

DB 120 SSHPHILVMLKGYBAPQIALRCGIMLRCEIRHEPLAKIILWSEQYDFFRYVEMSTFDI 179
QY 181 ASDAFATKOLLTRHKLLSABELEGHYDRFPFSEYEKLLHSENYVTKRQSLKLLGELLDR 240
DB 180 ASDAFATKOLLTRHKLLSABELEGHYDRFPFSEYEKLLHSENYVTKRQSLKLLGELLDR 239
QY 241 HNFITMTKYISKPENLKMMLLRDKSRNIOPEAFHFVKFVFNPNKTOPILDILLKNOA 300
DB 240 HNFITMTKYISKPENLKMMLLRDKSRNIOPEAFHFVKFVFNPNKTOPILDILLKNOA 299
QY 301 KLIEFLSFQNDRTDCMSSVPTTNSRVDLRVKPRTRGIRDILKRP 346
DB 300 KLIEFLSSFOKERTD-----DEQFADEKNYLIKQIRDLKKA 336

RESULT 14
Q803V8 ID Q803V8 PRELIMINARY; PRT; 343 AA.
AC Q803V8;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Zgc:55451.
GN Name=zgc:55451;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=AB; TISSUE=whole body;
RA MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RP SEQUENCE FROM N.A.
RA STRAIN=AB; TISSUE=whole body;
RA Strausberg R.;
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044172; AAH44172.1; -.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 343 AA; 39820 MW; F10450DA0446268A CRC64;

Query Match
Best Local Similarity 72.3%; Score 1287; DB 2; Length 343;
Matches 253; Conservative 42; Mismatches 41; Indels 18; Gaps 3;

QY 1 MPFPFGKSHKSPADIVKNLKSMAVLEKQDISDKKAEKATEEVSQNLVAMKEILYGTNE 59
DB 1 MPFPFGKSHKSPADIVKNLKSMAVLEKQDISDKKAEKATEEVSQNLVAMKEILYGTNE 60
QY 60 KEPTTEAVAQAQLYNSGLLSTLVADLQIDFEGKDVQVAFNNILRRQIGTRTPTVEYI 119

```

Db 297 EKLVDLTFHTDRSE 312

Search completed: October 22, 2004, 09:12:26
Job time : 197 secs

Db 61 KEPTAEVAQALQELYNLLISLIANLQRIDFEGKDVVHLFSNVRQIGARTPTVEY 120
QY 120 ICTQONILFMLKGYESPALNCGIMLRECIHREPLAKIILWSEOFYDFRYVEMSTFD 179
Db 121 ISHSQILFMLKGYETSEVALNCGMMLRECIHREPLAKIILWSEOFYDFRYVEMSTFD 180
QY 180 IASDAFATPKDLLTRHKLISABFLEOHYDRFFSEYEKLLHSENYVTKRQSLKLGELLDD 239
Db 181 IASDAFATPKDLLTRHKLISABFLEOHYDRFFSEYEKLLHSENYVTKRQSLKLGELLDD 240
QY 240 RHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQ 299
Db 241 RHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQ 300
QY 300 AKLIEFLSKFQNDRTD-----CKSSSVPTTNSRVDLRVKPRTRGIRDLKRPAPAE 349
Db 301 SKLVEFLSHFOTDRSEDFQDEKNY-----LIKQIRDLKRPAPAE 341

RESULT 15

Q7PF07 PRELIMINARY; PRT; 332 AA.
AC Q7PF07;
DT 01-MAR-2004 (T-EMBLrel. 26, Created)
DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE ENSANGP0000023148 (Fragment).
GN Name=ENSANGG0000021597;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAA01008804; EAA45510.1; -.
DR InterPro; IPR004892; MO25.
DR Pfam; PF03204; MO25; 1.
FT NON TER 332 332
SQ SEQUENCE 332 AA; 38665 MW; E315A6819F3BDA67 CRC64;

Query Match 65.3%; Score 1163.5; DB 2; Length 332;
Best local similarity 71.5%; Pred. No. 1.5e-66;
Matches 226; Conservative 46; Mismatches 39; Indels 5; Gaps 3;

QY 1 MPFPFGKSHKSPADIVKNLKSMAVLEKQDISDKKAEKATEEVSKNLVAMKEILYGTNEK 60
Db 1 MPL-FGSKSQSPQELVSLKREAVALER--GDKKAQKQEDVKNLVSKNMLCGTADA 56

QY 61 EPQTE-AVAQALQELYNLLISLIANLQRIDFEGKDVVHLFSNVRQIGARTPTVEY 119
Db 57 EPQTEVISQALHELSTDLILLIQLNLRIDFEGKDVVHLFSNVRQIGARTPTVEY 116

QY 120 ICTQONILFMLKGYESPALNCGIMLRECIHREPLAKIILWSEOFYDFRYVEMSTFD 179
Db 117 ICTKPEILFTLMAGYEHQIHALNCGIMLRECIHREPLAKIILWSEOFYDFRYVEMSTFD 176

QY 180 IASDAFATPKDLLTRHKLISABFLEOHYDRFFSEYEKLLHSENYVTKRQSLKLGELLDD 239
Db 177 IASDAFATPKDLLTRHKLISABFLEOHYDRFFSEYEKLLHSENYVTKRQSLKLGELLDD 236

QY 240 RHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQ 299
Db 237 RHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQ 296

QY 300 AKLIEFLSKFQNDRTD 315
Db 301 SKLVEFLSHFOTDRSEDFQDEKNY-----LIKQIRDLKRPAPAE 341

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2004, 09:12:34 ; Search time 132 Seconds
(without alignments)
858.451 Million cell updates/sec

Title: US-10-089-688-2
Perfect score: 1781
Sequence: 1 MPFFGKSHKSPADIVKNLK.....RVKPTRGIRDLKRPQQA 350

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1364641 seqs, 323758627 residues

Total number of hits satisfying chosen parameters: 1364641

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/ECTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1635.5	91.8	341	16	US-10-408-765A-911
2	1635.5	91.8	496	14	Sequence 911, App
3	1635.5	91.8	552	14	Sequence 5, Appli
4	1614.5	90.7	341	14	US-10-239-079-5
5	1297.5	72.9	337	14	US-10-025-730-3
6	1117	62.7	339	14	US-10-025-730-1
7	1054.5	59.2	377	14	US-10-025-730-4
8	750	42.1	446	15	US-10-025-730-5
9	747	41.9	336	16	US-10-425-114-52177
10	741	41.6	412	15	US-10-437-963-161075
11	721.5	40.5	339	15	US-10-424-599-224185
12	649	36.4	351	15	US-10-425-114-46025
13	639.5	35.9	337	16	US-10-437-963-183779

14	635.5	35.7	342	15	US-10-425-114-48872	Sequence 48872, A
15	593	33.3	334	16	US-10-437-963-146612	Sequence 146612, A
16	501	28.1	337	15	US-10-424-599-209662	Sequence 209662, A
17	400.5	22.5	197	16	US-10-767-701-37391	Sequence 37391, A
18	379.5	21.3	173	15	US-10-424-599-158029	Sequence 158029, A
19	353.5	19.8	140	16	US-10-767-701-42587	Sequence 42587, A
20	321	18.0	184	15	US-10-424-599-149595	Sequence 149595, A
21	286.5	16.1	294	16	US-10-437-963-119830	Sequence 119830, A
22	255	14.3	161	15	US-10-425-114-70029	Sequence 70029, A
23	209	11.7	135	11	US-09-864-408A-6108	Sequence 6108, Ap
24	168	9.4	83	15	US-10-425-114-44450	Sequence 44450, A
25	135.5	7.6	116	15	US-10-424-599-210348	Sequence 210348, A
26	126	7.1	1182	15	US-10-282-122A-53445	Sequence 53445, A
27	125.5	7.0	74	15	US-10-425-114-56749	Sequence 56749, A
28	116.5	6.5	888	10	US-09-893-519A-73	Sequence 73, Appli
29	113	6.3	2184	14	US-10-304-095-6	Sequence 6, Appli
30	111.5	6.3	70	15	US-10-424-599-210351	Sequence 210351, A
31	111.5	6.3	790	16	US-10-408-765A-2178	Sequence 2178, Ap
32	110.5	6.2	797	16	US-10-437-963-163096	Sequence 163096, A
33	109.5	6.1	367	15	US-10-282-122A-45776	Sequence 45776, A
34	109.5	6.1	564	15	US-10-282-122A-61170	Sequence 61170, A
35	109	6.1	807	9	US-09-847-046-2	Sequence 2, Appli
36	109	6.1	825	14	US-10-295-027-546	Sequence 546, App
37	109	6.1	858	14	US-10-367-978-35	Sequence 35, Appl
38	108.5	6.1	719	15	US-10-282-122A-54440	Sequence 54440, A
39	108	6.1	54	16	US-10-437-963-166440	Sequence 166440, A
40	105.5	5.9	63	15	US-10-424-599-230655	Sequence 230655, A
41	105.5	5.9	680	9	US-09-864-761-47959	Sequence 47959, A
42	105.5	5.9	820	14	US-10-029-386-32324	Sequence 32324, A
43	105.5	5.9	3878	14	US-10-080-608A-11	Sequence 11, Appl
44	105.5	5.9	3899	14	US-10-171-311-4	Sequence 4, Appli
45	105.5	5.9	3907	14	US-10-171-311-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

; Sequence 911, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Faby, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 911
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-408-765A-911

Query Match	91.8%	Score	1635.5	DB	16	Length	341
Best Local Similarity	94.0%	Pred. No.	1.5e+129				
Matches	329	Conservative	1	Mismatches	11	Indels	9
						Gaps	2
Qy	1	MPFFGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGNEX	60				
Db	1	MPFFGKSHKSPADIVKNLKESMAVLEKQDISDKAEKATEEVSKNLVAMKEILYGNEX	60				
Qy	61	EPQTEAVAQAQLAQLYNSGLLSTLVADQLIDFEGKDVQVAFNNILRRQIGTRTPTEYI	120				
Db	61	EPQTEAVAQAQLAQLYNSGLLSTLVADQLIDFEGKDVQVAFNNILRRQIGTRTPTEYI	120				

121	CTQONILFMLLKGYSEPIALNCGIMURECIRHEPIAKILLWSEQYDFRVYEMSTPDI	180
Qy		
121	CTQONILFMLLKGYSEPIALNCGIMURECIRHEPIAKILLWSEQYDFRVYEMSTPDI	180
Db		
181	ASDAFATPKDLLTRHKLLSABFLQHYDRFSEYEKLLHSENVVTKRQSLKLLGELLDR	240
Qy		
181	ASDAFATPKDLLTRHKLLSABFLQHYDRFSEYEKLLHSENVVTKRQSLKLLGELLDR	240
Db		
241	HNFTIMTKYISKPENLKLMMNLDRKSRNTQFAFHVFKVFVANPNKTOPILDLKNOA	300
Qy		
241	HNFTIMTKYISKPENLKLMMNLDRKSRNTQFAFHVFKVFVANPNKTOPILDLKNOA	300
Db		
301	KLIEFLSKFQNDRTDCWSSSVPTNSRVLDLRVKPRTGIRDLKRPQAQOE	350
Qy		
301	KLIEFLSKFONDTE-----DEQFNDEKTVIVKO-----IRDLKRPQAQOE	341
Db		

Query Match	91.8%;	Score 1635.5;	DB 14;	Length 496;
Best Local Similarity	94.0%;	Pred. No. 2.5e-129;		
Matches 329;	Conservative 1;	Mismatches 11;	Indels 9;	Gaps 2
Qy	1	MPFPFGSHKSPADIVKNLKESMAVLEKQDIDSDDKAEKATBEVSKNLVAMKEILYGTNEK	60	
Db	156	MPFPFGSHKSPADIVKNLKESMAVLEKQDIDSDDKAEKATBEVSKNLVAMKEILYGTNEK	215	
Qy	61	EPQTEAVQAQLAEIYNSGLLSTLVADLQIDIDPEGKQDVAQIFNNILRRQIGTGTPTVEYI	120	
Db	216	EPQTEAVQAQLAEIYNSGLLSTLVADLQIDIDPEGKQDVAQIFNNILRRQIGTGTPTVEYI	275	
Qy	121	CTQQNILFMLLKGYESPEIALNCGIMLRRECIRHEPLAKIILWSEQFYDFFRVEMSTPDI	180	
Db	276	CTQQNILFMLLKGYESPEIALNCGIMLRRECIRHEPLAKIILWSEQFYDFFRVEMSTPDI	335	
Qy	181	ASDAFATFKOLLTRHKLLSABFLQHYDRPFSEYEKLIHSENYVTKROSLKILGELLDR	240	
Db	336	ASDAFATFKOLLTRHKLLSABFLQHYDRPFSEYEKLIHSENYVTKROSLKILGELLDR	395	
Qy	241	HNFTIMTKYISKPENLKLMMNLLRDKSNIQEAFAHFVKFVFNANPKTQPIILDILLKNQA	300	
Db	396	HNFTIMTKYISKPENLKLMMNLLRDKSNIQEAFAHFVKFVFNANPKTQPIILDILLKNQA	455	
Qy	301	KLIBFLSKFONDRTDCMSSSVPTTNSRVDLRVKPRTRGIRDUKRPAQOEA	350	
Db	456	KLIBFLSKFONDRTDCMSSSVPTTNSRVDLRVKPRTRGIRDUKRPAQOEA	515	

```

; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: ANIC-BP1-ligand
; FILE REFERENCE: ANIC-BP-1-ligand
; CURRENT APPLICATION NUMBER: US/10/239,079
; CURRENT FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Lexa-ANIC-BP-1
; OTHER INFORMATION: fusion protein
; US-10-239-079-6

```

RESULT 3
US-10-239-079-6
; Sequence 6, Application US/10239079
; Publication No. US20030148446A1
; GENERAL INFORMATION:

QY 1 MPFPFGKSHKSPADIVKXNLESMAVLEKQDISDKKAEKATEEVSXNVLAMKEILYGTNEK 60
DB 1 MPFPFGKSHKSPADIVKXNLESMAVLEKQDISDKKAEKATEEVSXNVLAMKEILYGTNEK 60
QY 61 EPQTEAVAQLAQELVNSGLLSTLVADIQLIDFEGKDVQAIFNNILRRQIGTRTPTVEYI 120
DB 61 EPQTEAVAQLAQELVNSGLLSTLVADIQLIDFEGKDVQAIFNNILRRQIGTRTPTVEYI 120
QY 121 CTQQNILFMLKGYESPALNCGIMLRECRHREPLAKIILWSQFYDFFRVEMSTFDI 180
DB 121 CTQQNILFMLKGYESPALNCGIMLRECRHREPLAKIILWSQFYDFFRVEMSTFDI 180
QY 181 ASDAFATFKOLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKRQSLKLLGELLIDR 240
DB 181 ASDAFATFKOLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKRQSLKLLGELLIDR 240
QY 241 HNFNTMTKYISKPENLKLMMNLLRDKSRNIQFEAFHFVKVFNPNKTOPILDILLKNQA 300
DB 241 HNFNTMTKYISKPENLKLMMNLLRDKSRNIQFEAFHFVKVFNPNKTOPILDILLKNQA 300
QY 301 KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVKPRTRGIRDLKRPAAQBEA 350
DB 301 KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVKPRTRGIRDLKRPAAQBEA 350

RESULT 5

US-10-025-730-1

; Sequence 1, Application US/10025730
; Publication No. US20030045466A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; APPLICANT: Gorgone, Gina A.

; TITLE OF INVENTION: CALCIUM BINDING PROTEIN

; FILE REFERENCE: PF-0635 US

; CURRENT APPLICATION NUMBER: US/10/025,730

; CURRENT FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: US/09/190,965

; PRIOR FILING DATE: 1998-11-13

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PERL Program

; SEQ ID NO 1

; LENGTH: 337

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE: -

; OTHER INFORMATION: 3734805

US-10-025-730-1

Query Match

Best Local Similarity 72.9%; Score 1297.5; DB 14; Length 337;

Matches 263; Conservative 32; Mismatches 38; Indels 13; Gaps 3;

QY 1 MPFPFGKSHKSPADIVKXNLESMAVLEKQDISDKKAEKATEEVSXNVLAMKEILYGTNEK 60
DB 4 MPL-FSKSHKNPABIVKILKONLAILEKQ--DKTKDASEEVSXSLQAMKEILCGTNEK 59
QY 61 EPQTEAVAQLAQELVNSGLLSTLVADIQLIDFEGKDVQAIFNNILRRQIGTRTPTVEYI 120
DB 60 EPQTEAVAQLAQELVNSGLLSTLVADIQLIDFEGKDVQAIFNNILRRQIGTRTPTVEYI 119
QY 121 CTQQNILFMLKGYESPALNCGIMLRECRHREPLAKIILWSQFYDFFRVEMSTFDI 180
DB 120 SAHPHILFMLKGYEAQIALRCGIMLRECRHREPLAKIILWSQFYDFFRVEMSTFDI 179
QY 181 ASDAFATFKOLLTRHKLLSAEFLQHYDRFFSEYEKLLHSENYVTKRQSLKLLGELLIDR 240
DB 180 ASDAFATFKOLLTRHKLVADFLQNYDTIFEDYEKLLQSENYVTKRQSLKLLGELLIDR 239
QY 241 HNFNTMTKYISKPENLKLMMNLLRDKSRNIQFEAFHFVKVFNPNKTOPILDILLKNQA 300

DB 240 HNFNTMTKYISKPENLKLMMNLLRDKSRNIQFEAFHFVKVFNPNKTOPIVEILLKNQP 299
QY 301 KLIEFLSKFQNDRTDCMSSSVPTTNSRVDLRVKPRTRGIRDLKRPAA 346
DB 300 KLIEFLSSFOKERTD-----DEQFADEKNYLIQIRDLKJKA 336

RESULT 6

US-10-025-730-4

; Sequence 4, Application US/10025730

; Publication No. US20030045466A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; APPLICANT: Gorgone, Gina A.

; TITLE OF INVENTION: CALCIUM BINDING PROTEIN

; FILE REFERENCE: PF-0635 US

; CURRENT APPLICATION NUMBER: US/10/025,730

; CURRENT FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: US/09/190,965

; PRIOR FILING DATE: 1998-11-13

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PERL Program

; SEQ ID NO 4

; LENGTH: 339

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE: -

; OTHER INFORMATION: g1794137

US-10-025-730-4

Query Match 62.7%; Score 1117; DB 14; Length 339;

Best Local Similarity 69.0%; Pred. No. 8e-86; Matches 220; Conservative 51; Mismatches 40; Indels 8; Gaps 5;

QY 1 MPFPFGKSHKSPADIVKXNLESMAVLEKQDISDKKAEKATEEVSXNVLAMKEILYGTNEK 60
DB 1 MPL-FGKSQSPVELVKSLEAINALE---AGDRKVEKAQEDVSKNLVSIKNMLHGSSDA 56
QY 61 EPQTE-AVAQLAQELVNSGLLSTLVADIQLIDFEGKDVQAIFNNILRRQIGTRTPTVEY 119
DB 57 EPADYVVAQLSQBELVNSNLLLLIQNLHRIIDFEGKKHVALIFNNLLRRQIGTRTPTVEY 116
QY 120 ICTQONILEMLLKGYES--PRIALNCGIMLRECRHREPLAKIILWSQFYDFFRVEMST 177
DB 117 ICTAPEILFTLMAGYEDAHPEIALNSGMTURECARYEALAKIMLHSDDEFFFRYVEVST 176
QY 178 FDIASDAFATFKOLLTRHKLLSAEFLQHYDRFFSE--YEKLLHSENYVTKRQSLKLLGEL 236
DB 177 FDIASDAFATFKOLLTRHKLLCAEFLDANYDKFFSOHYQRLNSENYYVTRRQSLKLLGEL 236
QY 237 LLDHNTMTKYISKPENLKLMMNLLRDKSRNIQFEAFHFVKVFNPNKTOPILDILL 296
DB 237 LLDHNTMTKYISKPENLKLMMNLLRDKSRNIQFEAFHFVKVFNPNKTOPILDILL 296
QY 297 KNOAKLIEFLSKFQNDRTD 315
DB 297 RNQTKLVDFLTNFTNDRSE 315

RESULT 7

US-10-025-730-5

; Sequence 5, Application US/10025730

; Publication No. US20030045466A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; APPLICANT: Gorgone, Gina A.

; TITLE OF INVENTION: CALCIUM BINDING PROTEIN

; FILE REFERENCE: PF-0635 US

; CURRENT APPLICATION NUMBER: US/10/025,730

```

; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US/09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; OTHER INFORMATION: g1255838
; US-10-025-730-5

Query Match          59.2%; Score 1054.5; DB 14; Length 377;
Best Local Similarity 62.5%; Pred. No. 1.7e-80;
Matches 205; Conservative 49; Mismatches 61; Indels 13; Gaps 2;

QY 1 MPPFGKSHKSPADIVKNIKESMAVLEK-----QDISKKAETEEVSKNLYAM 50
Db 1 MLLFGKSHKSPADVVTTLREVLTILDKLPPPKDKDGNITQSDKKYDKALDEVSKNAMI 60
QY 51 KETLYGNEKEPOTE---AVAQAQELYNLGLSTLVADQLIDFEKGKDVQAFNNILR 107
Db 61 KSIYGNDSAPRSESHVQVAQAQEVYANILPMLIKMLPKFEFECKDVQGFNNLLR 120
QY 108 ROIGTRPTVEYICTQONILFMLKGYSPFIALNCGIMLRECIHREPLAKIILWSEQFY 167
Db 121 ROIGTRSPVEYLGARPEILLQLVQGSVVDIALTCGLMLRESIRHDLAKIILYSDVY 180
QY 168 DFRYVEMSTDIASDAFATFKOLLTRHKLLSAEFLQHYDRFFSEYKELHSENIVTKR 227
Db 181 TFEYVQSEVFDISSDAFSTFKELTTRHKAILAEFLDSNYDTFFAQYQNLNSKNYVTR 240
QY 228 QSLKLGELLDRHNFMTIMTKYISKPENLKMMLLRDKSRNIQFEAFHVKFVFNPNK 287
Db 241 QSLKLGELLDRHNFMTIMTKYISNPDLRMMLLDRKSRNIQFEAFHVKFVFNPNK 300
QY 288 TQPIDILLKNQAKLIFLSKFQNDRTD 315
Db 301 KPFIISILNENREKLVEFLSEFHNDRTD 328

RESULT 8
US-10-425-114-52177
; Sequence 52177, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 52177
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3912-007-H10_FLI.pep
; US-10-425-114-52177

Query Match          42.1%; Score 750; DB 15; Length 446;
Best Local Similarity 46.7%; Pred. No. 1e-54;
Matches 147; Conservative 64; Mismatches 100; Indels 4; Gaps 2;

QY 5 FGKSHKSPADIVKNIKESMAVLEKQDIS---DKKAETATEEVSKNLYAMKEILYGTNEKE 61
; US-10-425-114-52177
; Sequence 52177, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 52177
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3912-007-H10_FLI.pep
; US-10-425-114-52177

Query Match          41.9%; Score 747; DB 16; Length 336;
Best Local Similarity 46.3%; Pred. No. 1.3e-54;
Matches 146; Conservative 66; Mismatches 99; Indels 4; Gaps 2;

QY 5 FGKSHKSPADIVKNIKESMAVLEKQDIS---DKKAETATEEVSKNLYAMKEILYGTNEKE 61
Db 5 FKSXPRTPADVVRQTRRELLIFDLHSGRGGDAKREEMAEKSNIELKSLYNGESE 64
QY 62 POTEAVAQAQELYNLGLSTLVADQLIDFEKGKDVQAFNNILRQIGTRPTVEYIC 121
Db 65 PVTEACVQLTQEPFRENTRLLIICLPKMLNTRKDATQVVANIQRQVSSKIVASEYLE 124
QY 122 TQONILFMLKGYSPFIALNCGIMLRECIHREPLAKIILWSEQFYDFRIVEMSTFDIA 181
Db 125 ANKDLDTLISGYENMDIALHYGSMLECIHQSIARIVLESDHMKKFFDYIQLPNFDIA 184
QY 182 SDAPATFKOLLTRHKLLSAEFLQHYDRFFSEYK-ELLHSENIVTKRQSLKLGELLDR 240
Db 185 SDASATFKELLTRHKATVAEFLSNNDYDWEFFENSLSTSNVITTKRQAKILGDMMLDR 244
QY 241 HNFITMTKYISKPENLKMMLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQA 300
Db 245 SNSIVVMRYSSKDNLMILMLLRDSSKNITQIEAFHVKFLFAANKNKPPEVNVILVTRNS 304

```

```

Db 115 FKSXPRTPDVIVRQTRRECLVYLDLHSDSRGDAKREEMAEKSNIELKSLYNGESE 174
QY 62 POTEAVAQAQELYNLGLSTLVADQLIDFEKGKDVQAFNNILRQIGTRPTVEYIC 121
Db 175 PVTEACVQLTQEPFRENTRLLIICLPKMLNTRKDATQVVANIQRQVSSKIVASEYLE 234
QY 122 TQONILFMLKGYSPFIALNCGIMLRECIHREPLAKIILWSEQFYDFRIVEMSTFDIA 181
Db 235 SNKDLDTLISGYENMDIALHYGSMLECIHQSIARIVLESDHMKKFFDYIQLPNFDIA 294
QY 182 SDAPATFKOLLTRHKLLSAEFLQHYDRFFSEYK-ELLHSENIVTKRQSLKLGELLDR 240
Db 295 SDASATFKELLTRHKATVAEFLSNNDYDWEFFENSLSTSNVITTKRQAKILGDMMLDR 354
QY 241 HNFITMTKYISKPENLKMMLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQA 300
Db 355 SNAAVVMRYSSKDNLMILMLLRDSSKNITQIEAFHVKFLFAANKNKPPEVNVILVTRNS 414
QY 301 KLIEFLSKTQNDRTD 315
Db 415 KULRFFAGPKIDKBD 429

RESULT 9
US-10-437-963-161075
; Sequence 161075, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 161075
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_60293C.1.pep
; US-10-437-963-161075

Query Match          41.9%; Score 747; DB 16; Length 336;
Best Local Similarity 46.3%; Pred. No. 1.3e-54;
Matches 146; Conservative 66; Mismatches 99; Indels 4; Gaps 2;

QY 5 FGKSHKSPADIVKNIKESMAVLEKQDIS---DKKAETATEEVSKNLYAMKEILYGTNEKE 61
Db 5 FKSXPRTPADVVRQTRRELLIFDLHSGRGGDAKREEMAEKSNIELKSLYNGESE 64
QY 62 POTEAVAQAQELYNLGLSTLVADQLIDFEKGKDVQAFNNILRQIGTRPTVEYIC 121
Db 65 PVTEACVQLTQEPFRENTRLLIICLPKMLNTRKDATQVVANIQRQVSSKIVASEYLE 124
QY 122 TQONILFMLKGYSPFIALNCGIMLRECIHREPLAKIILWSEQFYDFRIVEMSTFDIA 181
Db 125 ANKDLDTLISGYENMDIALHYGSMLECIHQSIARIVLESDHMKKFFDYIQLPNFDIA 184
QY 182 SDAPATFKOLLTRHKLLSAEFLQHYDRFFSEYK-ELLHSENIVTKRQSLKLGELLDR 240
Db 185 SDASATFKELLTRHKATVAEFLSNNDYDWEFFENSLSTSNVITTKRQAKILGDMMLDR 244
QY 241 HNFITMTKYISKPENLKMMLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQA 300
Db 245 SNSIVVMRYSSKDNLMILMLLRDSSKNITQIEAFHVKFLFAANKNKPPEVNVILVTRNS 304

```

QY 301 KLIEFLSKFQNDRTD 315
||: ||: ||: ||:
Db 305 KLLRFFAGFKIDKED 319

RESULT 10

US-10-425-114-54669
; Sequence 54669, Application US/10425114
; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 54669

; LENGTH: 412

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3069-003-D7_FLI.pep

US-10-425-114-54669

Query Match 41.6%; Score 741; DB 15; Length 412;
Best Local Similarity 46.3%; Pred. No. 5.4e-54;
Matches 146; Conservative 64; Mismatches 101; Indels 4; Gaps 2;

QY 5 FGKSHKSPADIVKXNLSKESMAVLEKQDIS---DKKAEKATEEVSXKLVAMKILYGTNEKE 61

Db 81 FTKPTPTVDIVQTRCELVHLDHSGSRSGDKRDEKXTELSKRNIRDMKSILYNGESE 140

QY 62 PTEAVAQLAQELNSGLLSTLVADQLIDFEGKDVQAIFNNILRRQIGTRTPTVEYIC 121

Db 141 PVTEACVOLTOEFFRENTLRLLIHLPKLNETRKATQVAVNLRQOVSSKIVASEYLE 200

QY 122 TQOQNLFMLLKGYSPETALNCGIMLRCEIRHEPIAKIILMSEQYDFFRYVEMSTFDIA 181

Db 201 SNKDLIDLILGYENMDIALHYGAMLRCEIRHQSIARYVLESEHMKTFDVIQLPNFDIA 260

QY 182 SDATATPKDLTRKLLSABEFLQHYDRFFSEY-EKLIHSENYVTKROSLKILGELLILDR 240

Db 261 SDASATFKELTRKATVAEFLSNNDYDFFEFNSRLSSNTYITKROAIKLLGDMLLDR 320

QY 241 HNFITMTKYISKPENLKLMMNLLRDKSRNIQPEAFHVFVKFVAVNPNTQPTILDILLKNOA 300

Db 321 SNVAVMRYVSSKONLMTLMLNLLRDSKNIQIEAFHVFVKFAANKGKPEPVNVLITNRN 380

QY 301 KLIEFLSKFQNDRTD 315

Db 381 KLLRFFAGFKIDKED 395

RESULT 11

US-10-424-599-224185
; Sequence 224185, Application US/10424599
; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 224185

; LENGTH: 339

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_44468C.1.pep

US-10-424-599-224185

Query Match 40.5%; Score 721.5; DB 15; Length 339;

Best Local Similarity 48.5%; Pred. No. 1.8e-52;

Matches 150; Conservative 58; Mismatches 98; Indels 3; Gaps 2;

QY 1 MPFPFGKS--HKSPADIVKXNLSKESMAVLEKQDISDKXA-EKATEEVSXKLVAMKILYCT 57

Db 1 MSFSEFKALRPKTPQEVAKSIKESLMALDTKIVVEVKALEKALEVEKNFVTRMTLSGD 60

QY 58 NEKEPQTEAVAQLAQELNSGLLSTLVADQLIDFEGKDVQAIFNNILRRQIGTRTPTV 117

Db 61 GESEPNLDQVSQVVEICKEDVLTLLIHLKPLTGWEAKDLVHCWSILLKHKVEITNYCV 120

QY 118 EYICTQOQNLFMLLKGYSPETALNCGIMLRCEIRHEPIAKIILMSEQYDFFRYVEMST 177

Db 121 EYIEQHIELLDLVVVCYDNKDIALSCGIMLRCEIRKFPISLARYILESASFVLFKFEVLPN 180

QY 178 FDIASDAFATPKDLTRKLLSABEFLQHYDRFFSEY-EKLIHSENYVTKROSLKILGELL 237

Db 181 FDVASDAFSTPKDLTRKLVNVVSEFLTAYHDEFFDLIEKLTSPNYVTRQSLKLSFL 240

QY 238 LDRHNTMTKYISKPENLKLMMNLLRDKSRNIQPEAFHVFVKFVAVNPNTQPTILDILLK 297

Db 241 LESPSQIMKQYILEVRYLVKVMWTLTRDSSKNIQISAFHIFKVFVANPKPREVKIILSK 300

QY 298 NQAKLIEFL 306

Db 301 NQEKLLDLL 309

RESULT 12

US-10-425-114-46025

; Sequence 46025, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 46025

; LENGTH: 351

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: 700170954_FLI.pep

US-10-425-114-46025

Query Match 36.4%; Score 649; DB 15; Length 351;

Best Local Similarity 44.0%; Pred. No. 2.5e-46;

Matches 136; Conservative 62; Mismatches 101; Indels 10; Gaps 2;

QY 12 PADIVKXNLSKESMAVLEKQDISDKKAEKATEEVSXKLVAMKILYGTNEKEPQTEAVAQLA 71

Db 30 PQEATRSKQSLVALD-----TKGAKALEDAEKNIUTLRHTLAGDGEVPEPQSVQIA 84

QY 72 QELNSGLLSTLVADQLIDFEGKDVQAIFNNILRRQIGTRTPTVEYICTQQNILEWLL 131

Db 85 LEICKEGVLSLFQNLPSLWGAOKDLVHCWCILLRQKVDESVCQYIENHFDLDFLV 144
QY 132 KGYESPEIANGCMLRECIHREPLAKIILWSQFYDFRYYVEMSTFDIASDAFAFKDL 191
Db 145 VCYNLEVALNCGNMLRECIKYPTLAKYILESSFFLFQYVELSNFDFIASDALNTFKDL 204
QY 192 LTRHKLLSABFLQHYDRFSEYEKLLHSENYVTKRQSLKLLGELLIDRHNFTMTKYIS 251
Db 205 LTKHEDAVSEFLSHVEQFGLTKLLSSNYYTRQSVKFLSEFLLEAFNAQIMKRYIL 264
QY 252 KPENKLMNMLLRDKGRNIQFEAFHVKFVFNPNKTPQILDILLKNQAKLIEFL----- 306
Db 265 EVHYLNMGLLKDSSKNIRICAFHFKFVFNPNKPREIIQFIVENHREVLLHNLP 324
QY 307 SKFQNDRTD 315
Db 325 SKGEDEQLD 333

RESULT 13
US-10-437-963-183779
; Sequence 183779, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 183779
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(337)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_80837C.1.pap
US-10-437-963-183779

Query Match 35.9%; Score 639.5; DB 16; Length 337;
Best Local Similarity 40.4%; Pred. No. 1.5e-45;
Matches 135; Conservative 73; Mismatches 109; Indels 17; Gaps 2;
QY 11 SPADIVKNLKSMAVLEKQDISDKAEKATEEVSKNLVAMKEIILYGTNEKEPQTEVAQ 70
Db 15 SPQELVRSIKESLLALD-----TRTGAKALEDEVEKNVSTLRQTLGSGGEVEPNQEQVLQI 69
QY 71 AQELNSGLLSTLVADLQIDFEGKQDVAQIFNNILRQIGTRTPTEYICTQNNILFML 130
Db 70 ALEICKEDVLSLFVQNNPSLWGEGRKDLAHCWSILLRQKVDEAYCCVQYIENHFDLDFL 129
QY 131 LKGYESPEIANGCMLRECIHREPLAKIILWSQFYDFRYYVEMSTFDIASDAFAFKD 190
Db 130 VVCYNLEVALNCGNMLRECIKYPTLAKYILESSFFLFQYVELSNFDFIASDALNTFKD 189
QY 191 LTRHKLLSABFLQHYDRFSEYEKLLHSENYVTKRQSLKLLGELLIDRHNFTMTKYI 250
Db 190 LTRKHEAAVSFLCSHVEQFELYTRLLTSTNYYTRRQSVKFLSEFLLEAFNAQIMKRYI 249
QY 251 SKPENKLMNMLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQAKLIEFLSKFQ 310

Db 250 VEVSYLNIMIGLLKFSXKNIRICAFHFKFVFNPNKPRDIIQLVVDNHHLELLKL----- 305
QY 311 NDRDCMSSSVPTNSRVLDLRVKPRTRGRIDLKR 344
Db 306 -----GNLPTSKGEDQLEERDLIIKEIRK 331
RESULT 14
US-10-425-114-48872
; Sequence 48872, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 48872
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3637-244-EL_FLI.pap
US-10-425-114-48872

Query Match 35.7%; Score 635.5; DB 15; Length 342;
Best Local Similarity 43.0%; Pred. No. 3.3e-45;
Matches 131; Conservative 66; Mismatches 103; Indels 5; Gaps 1;
QY 11 SPADIVKNLKSMAVLEKQDISDKAEKATEEVSKNLVAMKEIILYGTNEKEPQTEVAQ 70
Db 25 TPEEVVRSIKDSFLALH-----TRTGAKALEDEVEKNVSTLRQTLGSGGEVEPNQEQVLQI 79
QY 71 AQELNSGLLSTLVADLQIDFEGKQDVAQIFNNILRQIGTRTPTEYICTQNNILFML 130
Db 80 TLEICKEDVLSLVQDLPSLWGEGRKDLAHCWCILLRQKVDEYCCVQYIENHLELDFL 139
QY 131 LKGYESPEIANGCMLRECIHREPLAKIILWSQFYDFRYYVEMSTFDIASDAFAFKD 190
Db 140 VGCYNLIDIALNCGNMLRECIKYPTLAKYILESSFFLFQYVELSNFDFIASDALNTFKD 199
QY 191 LTRHKLLSABFLQHYDRFSEYEKLLHSENYVTKRQSLKLLGELLIDRHNFTMTKYI 250
Db 200 LTRKHEVVAEFLSSHVEQFELYSRLLSSNYYTRRQSVKFLSEFLLEAFNAQIMKRYI 259
QY 251 SKPENKLMNMLLRDKSRNIQFEAFHVKFVFNPNKTPQILDILLKNQAKLIEFLSKFQ 310
Db 260 VEVRFINIMINLLKDSKNIRICAFHFKFVFNPNKPRCIVVALLDNRREVLLKHLNLP 319
QY 311 NDRTD 315
Db 320 SSKGD 324

RESULT 15
US-10-437-963-146612
; Sequence 146612, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad

```

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 146612
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4721C.1.pcp
US-10-437-963-146612

```

```

Query Match      33.3%; Score 593; DB 16; Length 334;
Best Local Similarity 39.2%; Pred. No. 1.2e-41;
Matches 125; Conservative 65; Mismatches 97; Indels 32; Gaps 4;

QY 5 FGKSHKSPADIVKNLKESMAVLEKQ-----DISDKKAEKATEEVSKNLVAMKEIILYGTNE 59
Db 23 FRSKARGPVEVVRHARELLAFLAENHDACGDKRDVKREHKMVDLAKSIGEMKSIILYNGE 82

QY 60 KEPQTEAVAQLAQELY--NSGLLSTILVADLQLIDFEGKKDVAQIFNNILRROIGTRPTV 117
Db 83 ADPVDEACSQLTKFEFFKENTSLHLVVCLPYMDLETQKDVQTANTLIG----- 132

QY 118 EYICTQQNILEFMLLKGYESPEIALNCGIMLRECIHSEPLAKIILWSEQFYDFFRYVEMST 177
Db 133 -----XRYDNMDIAIHYSAILRDCIRHQVAARYVLESQHKFFDIQYPD 178

QY 178 FDIASDAFATPKDLLTRHKLLSAEFLEQHYDRFFSEY-EKLLHSENYVTKRQSLKLGL 236
Db 179 ENIASDAFKTFKELLTRHSSAAEFSSNNYDWFPEFNSKLLSSNNYIIRQATQLGDI 238

QY 237 LLDRHNFIMTKYISKPENLKLMMNLLRDKSRNIQFEAFHVFVKVFPVNPNTQPIILDILL 296
Db 239 LLDKSNFTTAMVRYVSSKNLIIILMNLREQSRAIQVEAFRVFKLFTCNPKKPEIVGIIV 298

QY 297 KNOAKLIEFLSKFQNDRTD 315
Db 299 TNKSKILRFLADFTIEKD 317

```

Search completed: October 22, 2004, 09:24:17
Job time : 133 secs

